DETECTING TREND AND ABRUPT CHANGES IN MEAN AND EXTREME CLIMATE VARIABLES AND ASSESSING THE IMPACTS ON AGRICULTURE

BY

YAN GE

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Civil Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2019

Urbana, Illinois

Doctoral Committee:

Professor Ximing Cai, Chair
Professor Praveen Kumar
Professor Bo Li
Associate Professor Ryan Sriver
Dr. Tingju Zhu, The International Food Policy Research Institute

ABSTRACT

Climatic variables, temperature and precipitation, in particular, play critical roles in water resources management and agriculture management. To assess the impacts of climate change on human and natural environment, as well as to understand the causes of climate change, it is essential to investigate historic and current climate changes and predict potential future climate changes. This thesis is to develop a multiple change points detection method in time series to identify the change patterns of climate in the Continental US. The method will then be used to detect the changes in linear state changes and changes in variance slope with change time for multiple change points without knowing the number of change points before detection. It is found that abrupt changes occurred more frequently with precipitation than with temperature. Hot spots of identified changes in climate show closely correlated spatial-temporal patterns. The identified spatial-seasonal variation and correlation of changes highlights the uncertainty and information loss of averaging climate variables over years and the assumption of linear trends.

As global agricultural area has decreased recently, which may continue in the future, while population is increasing particularly in the developing counties, increased food production per unit area, i.e. yield, is required to meet growing food demand. Climate change is happening globally, with a general warmer trend and spatial difference in precipitation change. Particularly, more frequent extreme weather events (e.g., floods and droughts) have become a serious concern for agriculture, as well as human life and natural environment. Understanding how the change has affected crop yield to date will provide insights on its future possible impacts on food availability. The thesis will assess the impact of change in average monthly climate variables during the growing period on irrigated and rainfed crop yield to analyze whether impacts on yields are different during different months, also whether different impact on two types of yields.

In general, we find that the changes in mean climate variables during the past half century benefited less or hurt more the irrigated maize than the rainfed maize in Nebraska.

Besides mean change in climate, extreme climate change (e.g., drought and flood) may have changed or is changing. This thesis will focus on drought. Drought is a recurrent extreme climate phenomenon. It can last for weeks, months, even years, and the spatial extent of droughts is usually larger than other natural hazards (e.g., floods and hurricanes), resulting in devastating impacts on agriculture, water resources, environment and human lives. Meteorological drought events vary significantly from one region to another due to different climate characteristics. Additionally, internal variability in the Earth's climate causes the temporal variation of droughts, which has also been linked to climate change. Understanding the temporal and spatial characteristics of droughts can help in evaluating future drought risk and in choosing appropriate drought mitigation strategies. The thesis will investigate the spatial and temporal patterns of multiple drought characteristics (duration, severity and intensity) under different return periods in India and the Continental U.S (CONUS). In India, the temporal and spatial comparisons based on the univariate return period show different change patterns of duration, severity and peak intensity in different areas. Generally, in the areas which plant wheat more than rice, drought has been alleviated in duration and intensity after 1955; while in the areas which plant more rice than wheat, drought have been aggravated in duration, severity and intensity (except for area 8, a coastal area). In U.S., we find two significant patterns: Pattern I shows persistent droughts in Western & Eastern U.S., and the Great Plains, which experienced large variations in the drought characteristics over long time; Pattern II shows transient droughts in the interior of CONUS, which experienced short-term variations in drought characteristics; trends in these drought

characteristics at long and short return periods are different at some locations, showing the different trends of extreme and mild droughts.

ACKNOWLEDGEMENTS

Thanks to the excellent education, kindness and help that I received in Illinois since I came to U.S. in 2010.

First and foremost, I would like to express my most sincere thanks to my advisor, Prof. Ximing Cai. He has guided me all the way with consistent patience and great kindness. I appreciate all his contributions of time, ideas and funding to make Ph.D. experience fulfilling and productive. His attitude to job and passion to research has inspired me in the Ph.D. pursuit and will be an excellent example for my career. I also appreciate all the warmth that his family gives me like being home. It has been an honor and fortune to be his Master and Ph.D. student, and I am sincerely grateful to him and his family.

I would also like to thank the rest of my committee: Professor Praveen Kumar, Professor Bo Li, Professor Ryan Sriver, Dr. Tingju Zhu. In particular, Professor Kumar provided broader view of my research and suggestions of my research method, Professor Li offered solid guidance in statistics methods, Professor Sriver was generous with his time and ideas in helping me improve the results interpretation and giving suggestions of research, and Dr. Zhu kindly supported me with data collection and ideas for research. I heartily appreciate their great kindness of helping me to learn and improve.

I want to express my thanks to the funding support from the CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS) through the CCAFS project "Investigating the impacts of climate extremes on future water and food security" led by the International Food Policy Research Institute (IFPRI).

I also want to thank my parents for their unconditional love, support, devotion and giving me courage to chase my dream. I would like to thank my husband, Zhijian, who always supports

and lightens my days with his love when have being together for ten years. I am thankful to other members of my family for their support and love. I am grateful to friends who accompany me and warm me with friendship: Xiao Zhang, Zhenduo Zhu, Luo Hao, Liang Liu and many others. I also would like to thank my research group members: Tushar Apurv, Dr. Ruijie Zeng, Majid Shafieejood, Dr. Spencer Schnier, Dr. Yao Hu, Qianqun Zhao and all others.

TABLE OF CONTENTS

Chapter 1: Introduction

Climatic variables, temperature and precipitation, in particular, play critical roles in water resources management. The widely observed historical climate changes alter hydrologic cycle and add stress on water supplies and demands (Gleick et al., 2010). To assess the impacts of climate change on human and natural environment, as well as to understand the causes of climate change, it is essential to investigate historic and current climate changes and predict potential future climate changes, particularly at the regional scale (Hayhoe et al., 2007) so as to improve policy making and management methods adaptive to climate change. In particular, global agricultural area has decreased recently, which may continue in the future (Cassman et al., 2003; Young, 1999); population is increasing particularly in the developing counties. These situations require increased food production per unit area, i.e. yield, to meet growing food demand.

Climate change is happening globally, with a general warmer trend and spatial difference in precipitation change. Particularly, more frequent extreme weather events (e.g., floods and droughts) have become a serious concern for agriculture (Lobell et al., 2011), as well as human life (O'Hara and Georgakakos, 2008) and natural environment (Latta et al., 2010).

Understanding how the change has affected crop yield to date will provide insights on its future possible impacts on food availability (Lobell et al., 2011). More specifically, this thesis will focus on drought - both the impact of climate change on drought and the further impact of drought on crop yield. Drought is a recurrent extreme climate phenomenon. It can last for weeks, months, even years, and the spatial extent of droughts is usually larger than other natural hazards (e.g., floods and hurricanes) (Obasi, 1994), resulting in devastating impacts on agriculture, water resources, environment and human lives (Hao et al., 2014; WMO, 2006). Meteorological drought events vary significantly from one region to another due to different climate characteristics.

Additionally, internal variability in the Earth's climate causes the temporal variation of droughts, which has also been linked to climate change (Burke et al., 2006; Dai, 2013; Cook et al., 2015). Understanding the temporal and spatial characteristics of droughts can help in evaluating future drought risk and in choosing appropriate drought mitigation strategies.

The major research objectives included in this thesis are described as below.

To identify the change patterns of climate in the Continental US (Chapter 2)

Most previous studies analyzed climatic changes by simulating as a linear trend to demonstrate whether a linear trend exists in the given time series (Booth et al., 2011; Hayhoe et al., 2007; Pryor et al., 2009) in a chosen time range. Nonparametric test, Kendall's tau (Kendall, 1975) is commonly used to examine whether the given time series has significant linear trend (Booth et al., 2011; Pryor et al., 2009). When testing the trend rate of a given variable within a range, the time range is set and the test is not for detecting when the trend rate changes. Those studies examine whether there is a linear trend change but the timing of change and whether the changes exhibit other change patterns than a linear trend are not investigated.

In particular, abrupt changes that occurred frequently in the past may cause significant economic and ecological impacts (Alley et al., 2003). Regression method to detect gradual change cannot handle the abrupt change detection problem. Villarini et al. (2009) and Rouge et al. (2013) developed novel methods to detect and distinguish abrupt change pattern and gradual trend based on traditional nonparametric tests, which can provide a detected change pattern with occurring time. However, more complex change patterns besides abrupt change and gradual trend in a long-term time range (e.g. increasing trends with slope changed at one time) cannot be identified.

Detecting of past changes is essential to understand nonstationarity (Rouge et al., 2013), which is characterized by the change of time series in mean and/or variance. Previous studies usually only consider the systemic change in linear part, but not in the variance part, although the change of variance exists in real time series. Moreover, identifying occurrence time and patterns of changes is essential for cause analysis. As climatic variable has intrinsic system change itself, this situation also complicates detecting the changes caused by the change of internal physical dynamics or external influences. Change happening time and patterns may be related to various possible causes.

This part of the thesis will develop a multiple change points detection method in time series to identify the change patterns of climate in the Continental US. The method will then be used to detect the changes in linear state changes and changes in variance slope with change time for multiple change points without knowing the number of change points before detection. The detailed change patterns including change time of monthly, seasonal and annual precipitation, and maximum, average and minimum temperature are examined, while mostly focusing on monthly/seasonal ones.

Additional information about climate change detection can be found in Appendix A.

To estimate climate change impact on crop yield (Chapter 3)

The various impacts of climate change on crop yield will be analyzed by different climate scenarios. Most previous studies usually conducted the impact assessment as sensitivity analysis by generating climate scenarios via hypothesis, such as 2 degree warmer or 20% decrease of precipitation. However, as a start, it might be more necessary for us to understate how historical climate changes affected the crop yield. Lobell et al. (2011) simulated the climate change since 1980 as linear trends but more complex change patterns exist and have not been explicitly

considered. In this thesis study, average monthly climate variables during the growing period are processed to analyze whether impacts on yields are different during different months. In addition, separate irrigated yield and rainfed yield will enable the analysis of different impacts on the two yields from different climate variables.

Corn in United States is crucial to world food supply, accounting 41% of the world's corn Schlenker and Roberts, 2009). Thus, corn is chosen as a representative crop to be assessed for climate change impact. Statistical time series regression in county level in Nebraska and sensitivity analysis on climate scenarios generated according to change detection of historical climate with method described in Chapter 2, are used to assess the climate change impact on corn yield in the United States since 1947.

Additional information about climate change impact on crop yield can be found in Appendix B.

To identify drought pattern change under climate change (Chapter 4)

This chapter will extend the previous chapters to identify drought pattern change under climate change using Northern India Plain and the Continental U.S. as case studies

Drought occurs event by event; different events last a wide range of time periods and some can last over multiple years. Therefore, it is not appropriate to assess drought return periods using the classic procedures applied to intra-year extreme events (e.g., floods, Shiau and Shen, 2001), which only consider annual maximum. The length of a drought event and the cumulative severity can be more important than the maximum intensity of the event for assessing its impact on natural and human systems (e.g., agriculture and water demand). Thus, it has been suggested to use multiple characteristics including expected drought inter-arrival time, severity (S), duration (D), and peak intensity (I) for drought assessment (Shiau, 2006; Shiau and Shen,

2001). Given that the multiple drought characteristics are correlated, Copulas (Sklar, 1959) distributions are used to assess drought return periods based on those correlated variables. Understanding the change in multiple drought characteristics (D, S and I) is needed to examine both the separate and joint effect of the characteristics on agriculture.

One purpose of this thesis is to provide a more comprehensive drought assessment using PDSI at the regional level (i.e., northern and eastern India) underlying climate change. Compared to previous studies (Dai, 2013; Sheffield et al., 2012) that focus on trend analysis in time series of a drought index, our study distinguishes the changes in multiple drought characteristics. We use a time series of drought events to examine 1) the characteristics of the time series such as D, S and I of drought events, and 2) the changes associated with both individual and joint drought characteristics. Using the Copulas function to simulate the joint distribution of multiple drought variables, we will assess both the univariate and multivariate return periods of drought during three historical periods (1900-1954 and 1955-2012 of Dai (2013) and 1948-2008 of Sheffield et al. (2012)).

In India, drought occurs mostly due to the failure of south-west monsoon (June-September) (Reddy and Ganguli, 2012). About 33% of the arable land in India is considered to be drought-prone and a further 35% can also be affected under extreme climate conditions (Ganguli and Reddy, 2012). Northern and eastern India is a major agriculture area of India, especially for wheat and rice production. Drought frequency analysis in this region is needed to conduct risk evaluation and select drought-relief measures in this region. We will conduct the assessment in eight areas in our study region located in northern India plains. We will address the following questions: 1) has severe drought become more frequent, longer-lasting, or more intense over the historical period? 2) How are the two datasets (Dai and Sheffield et al.)

distinguished, in severe or mild droughts, or in long or short droughts? 3) Did the multivariate return period change over time and space?

The similar method investigating the spatial pattern of multiple drought characteristics in India is applied to the Continental U.S., while the method is extended to trend analysis of drought is based on the overlapping of moving windows proposed by Liu et al. (2014) to identify continuous changes in drought characteristics. In addition, drought patterns are identified whether they are persistent droughts (i.e., experiencing large variations in the drought characteristics over long time) or transient droughts (i.e., experiencing short-term variations in drought characteristics).

Finally, Chapter 5 presents conclusions and the description of future work.

1.1 References

Alley, R. B., et al. (2003), Abrupt climate change, *Science*, *299*(5615), 2005-2010.

Burke, E. J., Brown, S. J., Christidis, N. (2006), Modelling the Recent Evolution of Global Drought and Projections for the Twenty-First Century with the Hadley Centre Climate Model, *J. Hydrometeorol. 7*, 1113–1125.

Cassman, K., Dobermann, A., Walters, D. T., Yang, H. (2003), Meeting cereal demand while protecting natural resources and improving environmental quality, *Annual Review of Environmental Resources, 28*, 315-358.

Cook, B. I., Ault, T. R., Smerdon, J. E. (2015), Unprecedented 21st-century drought risk in the American Southwest and Central Plains, *Sci. Adv. 1*(1), e1400082, doi:10.1126/sciadv.1400082.

Dai, A. (2013), Increasing drought under global warming in observations and models, *Nat Clim Change, 3*(2), 171-171.

Ganguli, P., Reddy, M.J. (2012), Risk Assessment of Droughts in Gujarat Using Bivariate Copulas, *Water Resour. Manag. 26*(11), 3301-3327.

Gleick, P. H., et al. (2010), Climate Change and the Integrity of Science, *Science, 328*(5979), 689-690.

Hao, Z., AghaKouchak, A., Nakhjiri, N., Farahmand, A. (2014), Global integrated drought monitoring and prediction system, *Sci. Data 1*, 140001.

Hayhoe, K., et al. (2007), Past and future changes in climate and hydrological indicators in the US Northeast, *Clim Dynam, 28*(4), 381-407.

Kendall, M. G. (1975), *Rank Correlation Methods*, 2 ed., Griffin, London.

Latta, G., Temesgen, H., Adams, D., Barrett, T. (2010), Analysis of potential impacts of climate change on forests of the United States Pacific Northwest, *Forest Ecology and Management, 259*(4), 720-729.

Lobell, D. B., Schlenker, W., Costa-Roberts, J. (2011), Climate Trends and Global Crop Production Since 1980, *Science, 333*(6042), 616-620.

Obasi, G.O.P. (1994), WMO's role in the international decade for natural disaster reduction, *Bull. Am. Meteorol. Soc. 75*(9), 1655-1661.

O'Hara, J. K., and Georgakakos, K. P. (2008), Quantifying the Urban Water Supply Impacts of Climate Change, *Water Resources Management, 22*(10), 1477-1497.

Pryor, S. C., Howe, J. A., Kunkel, K. E. (2009), How spatially coherent and statistically robust are temporal changes in extreme precipitation in the contiguous USA? *International Journal of Climatology, 29*(1), 31-45.

Reddy, M.J., Ganguli, P. (2012), Application of copulas for derivation of drought severity–duration–frequency curves, *Hydrol. Process. 26*, 1672-1685.

Rougé, C., Ge, Y., Cai, X. (2013), Detecting gradual and abrupt changes in hydrological records, *Advances in Water Resources*, *53*(0), 33-44.

Sheffield, J., Wood, E. F., Roderick, M. L. (2012), Little change in global drought over the past 60 years, *Nature*, *491*(7424), 435.

Shiau, J. T. (2006), Fitting drought duration and severity with two-dimensional copulas, *Water Resour Manag*, *20*(5), 795-815.

Shiau, J. T., and Shen, H. W. (2001), Recurrence analysis of hydrologic droughts of differing severity, *J. Water Resour. Plann. Manage.*, *127*(1), 30–40.

Sklar, A. (1959), Fonctions de répartitions à n dimensions et leurs marges. *Publ. Inst. Stat. Univ. Paris*, *8*, 229–231.

Villarini, G., Serinaldi, F., Smith, J. A., Krajewski, W. F. (2009), On the stationarity of annual flood peaks in the continental United States during the 20th century, *Water Resources Research*, *45*(8).

WMO report (2006), Drought monitoring and early warning: concepts, progress and future challenges. WMO report no. 1006, World Meteorological Organization.

Young, A. (1999), Is there Really Spare Land? A Critique of Estimates of Available Cultivable Land in Developing Countries, *Environment, development and sustainability*, *1*(1), 3-18.

Chapter 2. Abrupt shifts and gradual trends of temperature and precipitation of the past century in the continental United States

2.1 Introduction

Changes in climate, such as more frequent extreme weather events (e.g., floods and droughts), has become a serious concern for human life (O'Hara and Georgakakos 2008), natural environment (Latta et al. 2010), and agriculture (Lobell et al. 2011). Numerous studies report changes in climate at the global, regional and even local scale (Hayhoe et al. 2007; IPCC 2013; Karl et al. 2009). However, we argue that some of previous estimates of changes are questionable in terms of whether they provide sufficient details about the changes in climate at the regional and local scale, as well as the methods used to obtain the results. Those estimates are based on a linear trend over a chosen time period (e.g., 15 or 50 or 100 years) with a usual assumption that changes, if existing, were gradual over time (IPCC 2013), implying a trend that is likely to continue. In fact, abrupt changes occurred frequently in the past and caused significant economic and ecological impacts (Alley et al. 2003). Abrupt changes refer to regime shifts from one condition to another (Lins and Slack 2005) within a relatively short time period at a rate triggered by a threshold process of the nonlinear climate system itself (Alley et al. 2002). Gradual trends and abrupt shifts require different adaptation measures, and making the distinction between the type of changes will help us understand the causal mechanisms in climatic changes at the global, regional, and local scale (e.g., (Villarini and Smith 2010; Villarini et al. 2009)). Moreover, if the changes exist, they can be documented with more certainty and being more convincing for policy making in determining necessary measures to mitigate or adapt to the existing changes or to those coming in the future (Lobell et al. 2011), by helping the policy and research communities identify when and where the changes occurred and at what scale such changes should be assessed for adaptation policies (i.e., global, regional, or local).

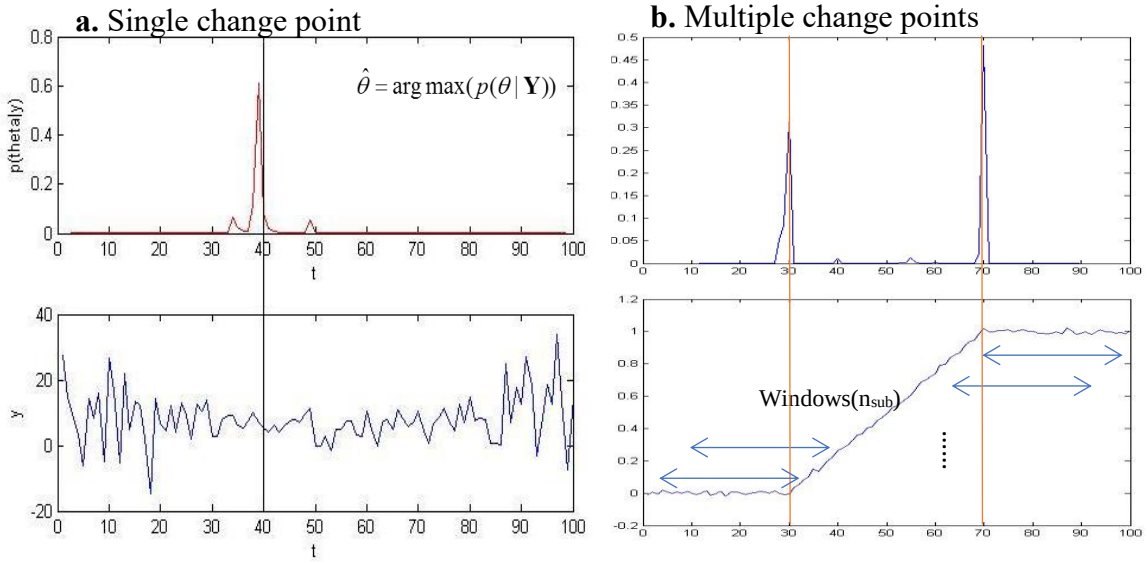
Usually, common-used distribution-free detection approaches (such as ranked-based methods, the Mann-Kendall test (Kendall 1975) and the Pettitt test (Pettitt 1979)) cannot detect the location of change points or the number of change points, which are however two essential goals for change point detection studies. Changes can be with the mean and/or variance of time series data; especially structure changes can occur with the relationship between a dependent variable (which can usually be observed) and independent (or explanatory) variables. The most common structure change is with the timing of the independent variable, i.e., when the change occurs in a given time series. Bayesian methods provide a measure of the uncertainty with an estimate of change, requiring distributional assumptions of observed time series, usually a normal distribution. They can deal with multiple types of change patterns and provide important information about the change process, such as the time of a change and number of change points. In our study, we improved upon a change point detection method for time series following Rougé et al. (2013) to detect the changes in linear states and changes in variant slopes with multiple change points in nonstationary climatic time series, without assuming the number of change points before detection. The method can then detect multiple mixed trends and abrupt shifts.

The method is applied to detecting detailed change patterns in temperature and precipitation of the past century in the continental United States. Important questions to answer include: are there any "hot spots" (areas dominated by one particular climatic change pattern) and where are they; how do the changes vary by season, and where are the ongoing trends? The study is based on the historical data from the United States Historical Climate Network (USHCN, 1910-2009) in the continental U.S.

2.2 Method

Bayesian frameworks have been applied to change point detection after the development of the change-in-mean method assuming normal observation by Chernoff and Zacks (1964) and assumption of normal or binomial observation by Smith (1975). To overcome the shortcoming of the early methods that only detect a single change point with a given time times, the Bayesian detection methods developed recently treat change points as abrupt shifts in mean or variance between two periods to deal with multiple change points(Chu and Zhao 2011; Hannart and Naveau 2009; Kim and Cheon 2010; Lai and Xing 2011; Lin et al. 2011; Rigaill et al. 2012; Tai et al. 2010). However, the Bayesian multiple regression approach (Fong and DeSarbo 2007; Schutz and Holschneider 2011) or Bayesian approach with time series models (e.g. autoregressive method) (Ray and Tsay 2002) considers the changes in the structure of regression components, often with an unknown number of change points, which can address the complexity of more realistic cases. Such Bayesian methods, including the Bayesian multiple regression approach (Fong and DeSarbo 2007; Schutz and Holschneider 2011) and the Bayesian approach with a time series model, allow for flexible relationships or models in different time periods so that the model formulas used in structure change detection can be various. In particular, Schutz and Holschneider (2011) used Bayesian inference with a piecewise-linear model to detect the linear trend changes of both mean and variance in time series. In this study we propose to combine the Pettitt test (Pettitt 1979; Rougé et al. 2013) and the Bayesian inference detection method developed by Schutz and Holschneider (2011). The latter provides the posterior probability density of a single change point location of a time series (Figure 2.1a) to detect the changes in linear states and in variant slopes with multiple change points without assuming the number of change points before detection. The Pettitt test is used to justify whether a change

point exists in the chosen partial time series (defined as a window) (Pettitt 1979; Rougé et al. 2013) to overcome expensive computation with the Bayes factor used in Schutz and Holschneider (2011). The combination of the two techniques enables the detection of multiple change points (Figure 2.1b). In addition, a novel procedure is added to the proposed method to justify which change points are statistically significant based on their final posterior probability density, and thus the number of change points can be determined.



* The final posterior density function of change points is computed as:

$$\sum_{windows} p(\theta \mid \text{cp in window}) p(\text{cp exists in window})$$, where cp denotes a change point.

Figure 2.1 Example figures of posterior density function of change point(s) (above) given a time series (below) with (**a**) one change point and (**b**) multiple change points.

The model in this study can consider the complexity of each regime (i.e., a segment of a time series before or after an abrupt change) either as a constant with noise or as a linear trend with noise based on changing variance, and thus the model will better simulate the complexity of hydroclimatic data time series than a piecewise constant model (Rougé et al. 2013) in representing abrupt changes between two segments . This method is also different from those

methods that assume a climatic variable time series follows a linear trend. With that assumption, the changes in climate at locations where abrupt changes occur would be misidentified as either no change (i.e., upward and downward abrupt changes can be cutoff) or as a rough linear trend. In addition, the information about the starting and ending point of a trend and the duration cannot be identified by simulating the time series as a linear trend. This is because the starting time and ending time of a trend are spatially distinct, and a simple linear simulation over a given time range will not be able to identify the spatial distinction.

Several typical types of artificial time series representing common change patterns are generated to validate the proposed local posterior density method for unknown multiple change points. Through the validation, the detection method is confirmed as an effective method for application to hydroclimatic time series. The details of the procedures are provided in the following.

a. Detection methods for a single change point

Following the work of Schutz and Holschneider (2011) based on Bayesian inference to detect trend changes in time series, a linear mixed model is formed to represent two aspects of change-points in time series: 1) a sudden change of the local linear trend and the change of slope and/or intercept; 2) a sudden change of the local variance. We assume that the change-points reflect the two aspects coincide in time, and the intercepts and slopes of the two linear-trend states are different; however, we do not assume the two continuous state intersect at the change point. Thus a time series with a single change point at time θ can be represented by a linear mixed model formulated as follows:

$$y(t) = \alpha_1(\tau_-^\theta(t)) + \beta_1(\varsigma_-^\theta(t)) + \alpha_2(\tau_+^\theta(t)) + \beta_2(\varsigma_+^\theta(t)) + \xi(t), \quad (1)$$

where functions $\varsigma_{-}^{\theta}(\varsigma_{+}^{\theta})$ and $\tau_{-}^{\theta}(\tau_{+}^{\theta})$ are defined as below to differentiate the slopes (β_1 and β_2) and intercepts (α_1 and α_2) in the pre-change state and the post-change state, respectively.

$$\varsigma_{-}^{\theta}=\begin{cases}\theta-t, \text{ if } t<\theta\\ 0 \quad , \text{else}\end{cases}, \varsigma_{+}^{\theta}=\begin{cases}t-\theta, \text{ if } t>\theta\\ 0 \quad , \text{else}\end{cases}$$
$$\tau_{-}^{\theta}=\begin{cases}1, \text{ if } t<\theta\\ 0, \text{ else}\end{cases}, \tau_{+}^{\theta}=\begin{cases}1, \text{ if } t>\theta\\ 0, \text{ else}\end{cases} \quad . \quad (2)$$

Fluctuations around the linear section of the model are expressed as ξ, which follows a Gaussian distribution with a zero mean and a variance which changes with time but remains constant at the change point time θ. We assume that this 'noise' is uncorrelated at different time points. The fluctuation sections of the model are defined as:

$$STD(\xi(t))=\sigma(1+s_1(\varsigma_{-}^{\theta}(t))+s_2(\varsigma_{+}^{\theta}(t))), \quad (3)$$

and

$$E(\xi(t)\xi(t'))=0, \ t\neq t' . \quad (4)$$

where σ is the scale factor describing the variability around the change point and constants $s_{1,2}$ reflects how the variance changes with time in the pre-change and post-change states.

The estimation of parameters in the model (change point time, linear coefficients, noise slope parameters and standard deviation scale factor) is based on a) Bayesian inference, b) the assumption of Gaussian distribution of observations as described above, and c) given simple prior distribution of parameters. The details of the parameter estimation should be referred to the work of Schutz and Holschneider (2011). The posterior distribution $p(\boldsymbol{\beta},\sigma,\mathbf{s},\theta\,|\,\mathbf{y})$ of parameters given observations is computed based on the process of Bayesian inference with given prior distributions. By integration, the marginal conditional distribution of each parameter can be obtained. Since the purpose here is to detect the change point time, our main goal is to integrate

the linear coefficients $\boldsymbol{\beta}$, noise slopes $\mathbf{s}$ and standard deviation scale factor σ in $p(\boldsymbol{\beta},\sigma,\mathbf{s},\theta \mid \mathbf{y})$ to obtain $p(\theta \mid \mathbf{y})$.

In order to decrease the impact of noise on change point detection, we use a local sum of the probability density function (Eq. (5)) to estimate the change point time by finding the interval (of length=5) with maximum p_{sum}:

$$p_{sum}(\theta \mid \mathbf{y}) = sum(p(\theta-2:\theta+2 \mid \mathbf{y})). \qquad (5)$$

We then estimate the change point time by computing the mathematical expectation in that interval defined as

$$\begin{aligned}\hat{\theta} &= \mathrm{E}(\arg\max(p_{sum}(\theta \mid \mathbf{y}))-2:\arg\max(p_{sum}(\theta \mid \mathbf{y}))+2) \\ &= \sum_{t=\arg\max(p_{sum}(\theta \mid \mathbf{y}))-2}^{\arg\max(p_{sum}(\theta \mid \mathbf{y}))+2} p(t \mid \mathbf{y}) \cdot t\end{aligned}. \qquad (6)$$

A parameter (p_{stop}) is used to judge whether the estimated change point time is an acceptable one: when $p_{sum} > p_{stop}$, we accept the estimated change point as a significant change point; otherwise, the estimated change is identified as insignificant. Other parameters such as linear coefficients, noise slope parameters and standard deviation scale factor can also be estimated through the marginal conditional distribution of each parameter; the value of a parameter is chosen when the marginal conditional distribution of the parameter is maximized.

b. Local posterior density method for unknown multiple change points

Two or more change-points can occur in a time series, especially for data collected over a long period of time. One challenge is there is no knowledge about how many change points exist before the detection of a particular change. Regressions of linear mixed models with a fixed number of change points hence require some comparison criteria among models with different

numbers of change points. Here, we propose to use local posterior density methods to detect multiple change points without prior knowledge.

Around each time point $t_j, j = 1, 2, ..., m(m < n)$, we choose a data window $I_j = [t_j - \frac{n_{sub}}{2} + 1, t_j + \frac{n_{sub}}{2}]$ with a length of n_{sub}. We assume there is at most one change point in one window. Thus by applying the detection method for a single change point to each window, we can compute the local posterior $p_{t_j}(\theta | y_{|I_j})$, which provides the posterior probability density of a possible change point within the window. To compute the global posterior distribution of multiple change points, the credibility on whether there is a change point within a window is needed as a weight for each local posterior (i.e., with each of the windows or sub-series). This is computed as:

$$\sum_{windows} p(\theta \,|\, \text{cp in window}) p(\text{cp exists in window}) \quad , \qquad (7)$$

where cp denotes a change point.

The window credibility is based on the Pettitt method (Pettitt 1979; Rougé et al. 2013), which is a non-parametric rank-based change point test. Its null hypothesis (H_0) is that when arbitrarily choosing a date, there is no change of the median between pre-date and post-date data. The test computes Pettitt statistics $k(\eta)$ to compare the rank before and after a date (η) for observations $y_i, i = 1, 2, ..., n$, by the following equation:

$$k(\eta) = \sum_{i=1}^{\eta} \sum_{j=\eta+1}^{n} \text{sgn}(y_j - y_i), \qquad (8)$$

where $\text{sgn}(y_j - y_i)$ is the sign of the difference of two different observations (y_j, y_i), as defined as:

$$\mathrm{sgn}(y_j - y_i) = \begin{cases} 1, y_j > y_i \\ 0, y_j = y_i \\ -1, y_j < y_i \end{cases}. \quad (9)$$

When absolute $k(\eta)$ takes the largest value, the change most likely occurs at time η. Ultimately, Pettitt statistics is K, which can be represented by:

$$K = \max_{\eta=1,2,..n} |k(\eta)|. \quad (10)$$

The null hypothesis rejection significance probability is approximated by:

$$p \approx 2\exp\left(\frac{-6K^2}{n^3 + n^2}\right). \quad (11)$$

For a significance level, α, the null hypothesis is rejected if $\alpha > p$. When the rejection significance probability (p) is relatively small, it is more likely to reject the null hypothesis, which states that there is more chance for a change existing in the time series. Thus $exp(-p)$ is defined as a weight for the window $p(\text{cp exists in window})$. In order to improve the change existence credibility, we define a parameter (p_{cut}) that judges whether there is a change point within the window. When $p > p_{cut}$, we set $p(\text{cp exists in window}) = 0$, i.e., there is no change point in the window; otherwise, $p(\text{cp exists in window}) = exp(-p)$.

In this study, a linear trend is considered as a linear non-zero state without a change point in a given sub-period (i.e., part of the study time series). For the observation samples used as input for a Pettitt test, the removal of linear trends, if existing, should be conducted first using a simple linear regression. This pretreatment can increase the accuracy of detecting change points between two trends as the requirement of Pettitt test (Rougé et al. 2013).

With the weights based on Pettitt test, the global posterior distribution of change point(s) can be obtained. Further, by the significance judgment (same as the one for the single change

point method), the time point(s) with the estimated significant change can be acquired. Following that, other parameters (including linear coefficients, noise slope parameters and standard deviation scale factor) can be estimated through the marginal conditional distribution of each parameter (same as the one for the single change point method).

c. Classification of change patterns

Through the local posterior density method for unknown multiple change points, the number and locations of transition points in a given time series can be estimated. Thus for i^{th} state $\mathbf{y}_{|(t_i, t_{i+1})}$ from time t_i to t_{i+1}, the noise slope $\mathbf{s}$ can estimated based on the posterior density $p(\mathbf{s} \mid \mathbf{y}_{|(t_i, t_{i+1})})$. With noise slope $\mathbf{s}$, the linear coefficients $\boldsymbol{\beta}$ and standard deviation scale factor σ can be estimated based on their posterior density distributions, respectively.

However, the slope of the linear trend slope in the linear coefficients β for one state must be tested for significance (5%) to justify whether it is a linear increase/decrease trend or a zero-slope state. Thus for each state, the change can be distinguished by an increasing trend, decreasing trend or flat state (i.e., no-change).

Therefore, given the estimated number and location of change point(s), as well as the pattern of each state based on the trend significance test (i.e., increasing trend, decreasing trend or flat state, as stated above), all change patterns with the entire time series can be identified (Figure 2.2). Note that a transition point can exist between three different states (zero-slope state, linear increasing trend, and linear decreasing trend), the various change patterns within any windows (i.e., segments of a time series) can be represented as a number of forms, as described in Figure 2.2. Especially, whether a change is an abrupt change (i.e., changing from one zero-slope state to another zero-slope state), monotonically increasing or decreasing trend, can be distinguished.

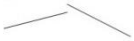
Number of change points	Schematic diagram(Simplified actual time series omitting noise)		Label	Description
0			Type I: No change	No change
			Type II: Increasing trend	Increasing trend
			Type III: Decreasing trend	Decrease trend
1			Type IV.I/D: Abrupt change. Increase/Decrease	Zero-slope state abruptly increases/decreases to another zero-slope state
			Type V.I/D: Increasing trend in later years. Increase/Decrease	Zero-slope state abruptly increases/decreases to a linear increasing trend
			Type VI.I/D: Decreasing trend in later years. Increase/Decrease	Zero-slope state abruptly increases/decreases to a linear decreasing trend
			Type VII.I/D: Increasing trend in former years. Increase/Decrease	A linear increasing trend abruptly increases/decreases to zero-slope state
			Type VIII.I/D: Slope-change increasing trends. Increase/Decrease	A linear increasing trend abruptly increases/decreases to another linear increasing trend
			Type IX.I/D: Decreasing trend in former years. Increase/Decrease	A linear increasing trend abruptly increases/decreases to a linear decreasing trend
			Type X.I/D: Decreasing trend in former years. Increase/Decrease	A linear decreasing trend abruptly increases/decreases to zero-slope state
			Type XI.I/D: Decreasing trend in former years. Increase/Decrease	A linear decreasing trend abruptly increases/decreases to a linear increasing trend
			Type XII.I/D: Decreasing trend in former years. Increase/Decrease	One linear decreasing trend abruptly increases/decreases to another linear decreasing trend
2			Type XIII.II/ID/DI/DD: Two abrupt changes. Two increases/ Increase. Decrease/ Decrease. Increase/Two decreases	Zero-slope state abruptly increases/decreases to another zero-slope state, and then abruptly increases/decreases to a third zero-slope state
				
			Type XIV.II/ID/DI/DD: An abrupt change, and an increasing trend. Two increases/ Increase. Decrease/ Decrease. Increase/Two decreases	Zero-slope state abruptly increases/decreases to another zero-slope state, then abruptly increases/decreases to a linear increasing trend
				

Figure 2.2 A schematic diagram of major change patterns

The method provides the posterior distribution of the change point(s). Based on a preset value of probability density according to tests, it can choose the significant changes. Then we justify whether the regimes before and after the detected change point has a significant (5% level) non-zero slope. Therefore, to determine a change pattern takes several steps, rather than providing a simple significance value of the test.

d. Validation

For the procedure based on the local posterior density to detect multiple change points, there are two important parameters p_{stop} and p_{cut} to estimate. p_{stop} is a parameter used to justify whether a given time is an acceptable change point; p_{cut} is a parameter used to justify whether there is change point within a window. A small value for p_{stop} may overestimate change points; while a large value of p_{stop} may skip some real change points. On the other hand, when p_{cut} is too small, some windows may be misjudged as ones that contain no change points; meanwhile some windows that do not contain any significant change point may be misjudged as ones that contain the change points. We generate some typical types of time series (e.g., abrupt change with changing variance; slope-change increase; two abrupt changes, etc.) to test the sensitivity of these two parameters to validate the proposed method (Ge 2012). The key procedure to choose proper values of these two parameters is to find the balance between detecting all change points and omitting nonexistent points. Through the trial-and-error validation process, $p_{stop} = 0.2$ and $p_{cut} = 0.6$ are chosen for all change point identifications conducted in this study.

The method was applied to the data sets of seasonal total precipitation, maximum and minimum daily temperature, and the diurnal temperature range (DTR, the difference between the daily maximum and minimum temperature) from 1910 to 2009. The data was acquired from USHCN with 1,218 stations across the Continental U.S. (Menne et al. 2012), which provides

sufficiently long (a century long) and spatially detailed (covering the whole continental US) enough to use our method to identify more details on the changes that occurred in the past and are ongoing at present. The classified change patterns for the four types of climatic variables listed above vary by season and by location (Figures A.1 to A.4 in Appendix A). The magnitude of the change (Figure A.5 in Appendix A) is defined to express how strong an abrupt change is; the gradient (slope) of the change (Figure A.6 in Appendix A) is defined to express how quick a gradual change is.

2.3 Results

a. Spatial-seasonal variability

The changes in climate experienced different spatial patterns in different seasons, consistent with the findings of previous work (Groisman et al. 2004; Hamlet 2011; Hayhoe et al. 2007) (Table 2.1). Comparing one particular change pattern over four seasons, one can find the spatial distribution of the change pattern differs considerably in each season. Figure 2.3 illustrates the spatial distribution of the stations that have linearly increasing trends for 100 years from 1910 to 2009 with the seasonal mean daily maximum temperature and seasonal total precipitation.

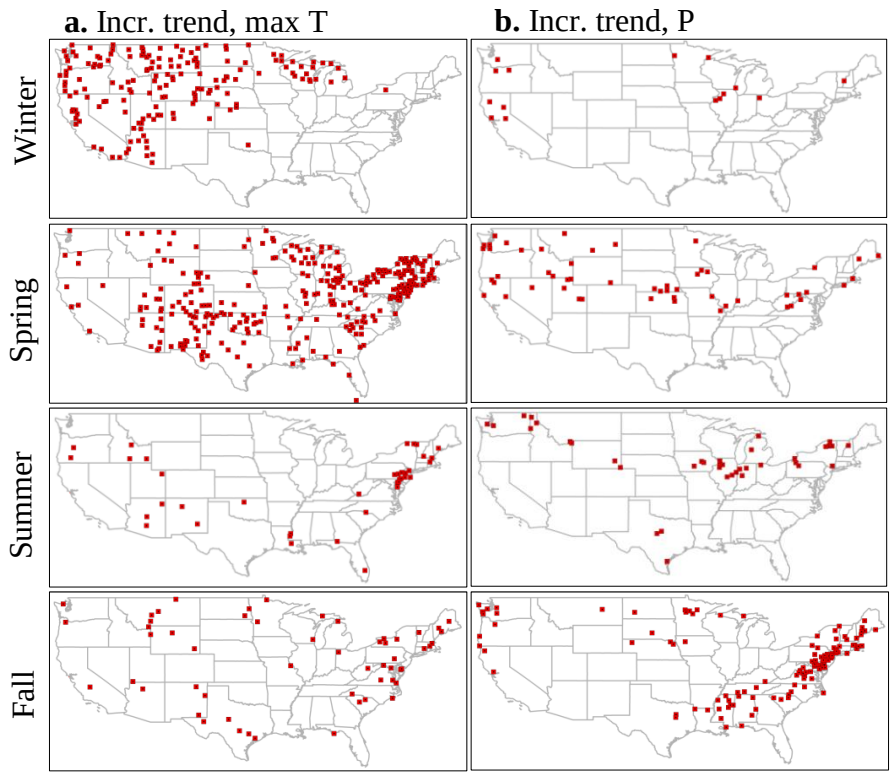


Figure 2.3 Spatial distribution of sites with 100-year increasing trend from 1910 to 2009 with (**a**) Seasonal mean maximum temperature and (**b**) Seasonal total precipitation.

In winter, the stations of the increasing trend for mean maximum temperature are mostly distributed in the western half of the continental U.S., especially the northwestern quadrant, while such stations spread to the entire continent (particularly the Northeast) in other three seasons, specifically in spring (Figure 2.3a). A study using the same dataset but from 1950 to 2002, indicated the increase trends in minimum temperature were located more in the northwestern quadrant of the country in both winter and spring (Groisman et al. 2004), while our study based on the dataset from 1910 to 2009 find that the continuing 100-year trend is located mainly in winter in the northwestern quadrant, but in spring in the eastern half of the continental U.S. It is found there were abundant but not all stations with abrupt downward or upward changes in northwestern country in the spring in 1950s or 1960s, and 1970s with increasing trend afterwards. This again reveals the information loss by pure linear trend detection. The

distributions of continuously increasing trends in total precipitation also reveal strong seasonal variability, although the number of stations with increasing trends of precipitation is much less than those for increasing temperature trends. Previous studies indicate increasing precipitation in winter and decreasing precipitation in summer in the Pacific Northwest region of North America (Hamlet, 2011), but opposite trends in the Northeast region (Hayhoe et al. 2007). However, our method revealed the increasing trend of seasonal total precipitation occurred in the North in winter and spring but such changes move south in the summer and fall (Figure 2.3b).

Table 2.1 Percentages of major change patterns for total seasonal precipitation, mean seasonal maximum daily temperature, mean seasonal minimum daily temperature and mean seasonal daily Diurnal Temperature Range (DTR) from 1910 to 2009

	Major change patterns		Winter (%)		Spring (%)		Summer (%)		Fall (%)	
Precipitation	No change		14.3		20.5		16.3		13.7	
	Increasing trend		1.2		4.4		3.1		9.6	
	Decreasing trend		1.5		0.4		0.8		0.7	
	Step change	Increase	46.6	26.2	45.4	30.7	45.6	29.2	44.8	32.4
		Decrease		20.4		14.7		16.4		12.4
	Gradual change(s) in part of the period		28.1		22.3		26.4		23.3	
Maximum temperature	No change		10.9		25.8		2.7		18.6	
	Increasing trend		13.7		24.8		3.1		4.6	
	Decreasing trend		0.3		2.3		1.1		3.6	
	Step change	Increase	35	23.9	24.8	21.3	32.9	19.1	33.9	20.6
		Decrease		11.2		3.5		13.8		13.3
	Gradual change(s) in part of the period		35.8		19		49.9		32.9	
Minimum temperature	No change		6.1		14.4		1.7		14.6	
	Increasing trend		16.4		23.2		3.9		11.8	
	Decreasing trend		0.3		1.8		0.1		1.2	
	Step change	Increase	35.5	26.6	21	16.5	26.5	22.8	27.5	18.5
		Decrease		8.9		4.6		3.7		9
	Gradual change(s) in part of the period		32		33.1		55.4		36.8	
DTR	No change		8.2		12.4		2.7		3.3	
	Increasing trend		3.7		7.2		1.2		1.7	
	Decreasing trend		5.8		9.2		5.8		3.6	
	Step change	Increase	22.9	10.2	25.6	12.8	19.2	6.1	31.7	10.9
		Decrease		12.7		12.7		13.1		20.9
	Gradual change(s) in part of the period		50.5		38.7		63.3		51.1	

b. Spatial-temporal correlations

Furthermore, we found a specific change type in climatic variables is jointly related with space and time (i.e., the occurrence year of a particular type of change). Spatially correlated changes in climate were identified and related to various factors, such as terrain (e.g., mountain or plain) and geographic location (e.g., coastal or inland, high latitude or low latitude) by previous studies (Alfaro et al. 2006; Allard and Keim 2007; IPCC 2013). However, the temporal correlation and the joint spatial and temporal correlation have rarely been discussed due to the method limitation in detecting a change time point. In our study, the magnitudes of an abrupt upward/downward change (shown by the sizes of symbols) in different occurrence years (shown by colors) present a clear spatial correlation (Figure 2.4). The occurrence years of abrupt upward changes in the summer maximum temperature appeared earlier in the north and later in the south (Figure 2.4a). The abrupt changes of summer maximum temperature, no matter downward or upward (Figure 2.4a and Figure A.1 in Appendix A) in the Midwest in 1930s are probably due to the 1936 heat wave in North America, a most severe one in modern history of North America, particularly in the Midwest(Andresen et al. 2012; Bridger et al. 1976).

In the Midwest, abrupt upward changes in the summer minimum temperature occurred in the early years of the study (e.g., 1920s and 1930s), while the changes occurred in later years in other regions (Figure 2.4b). The magnitudes of the abrupt changes in fall total precipitation in the western states are much smaller than those in the eastern half of the U.S. The abrupt upward change in fall total precipitation occurred across the East Coast mostly during the 1940s and 1950s. In addition, clear clusters of recent abrupt upward changes can be seen in locations from southwest Michigan to Indiana and down to southern Illinois (Figure 2.4c). The magnitudes of abrupt downward changes in the fall DTR are larger in the South, Northeast, and Southwest. In

the South and Southwest, the fall DTR decreased more than 1°C in the 1950s in a number of stations, while in the Northeast region, Ohio, and Indiana, the fall DTR decreased in the 1960s (Figure 2.4d).

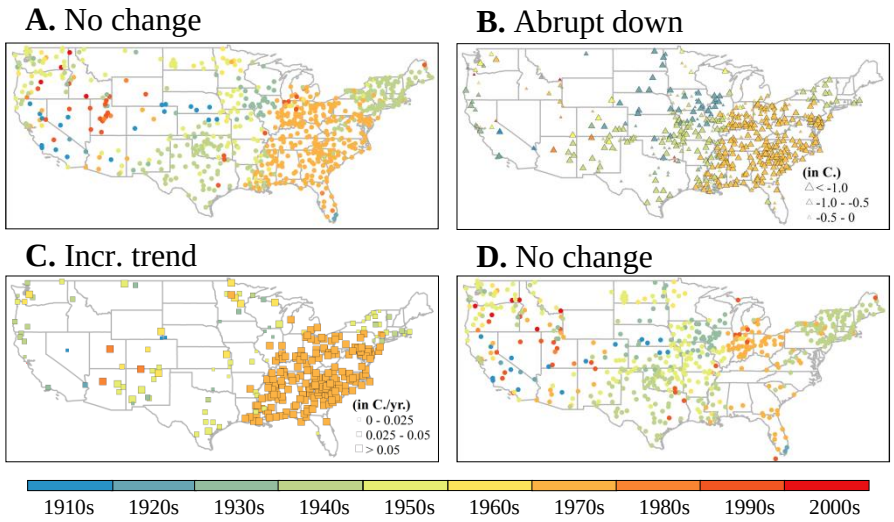


Figure 2.4 Changes in winter maximum temperature. (**A**) Spatial distribution of "No change" before occurrence year; (**B**) Magnitude of an abrupt downward change in an occurrence year; (**C**) Gradient of a gradually increasing trend after an occurrence year; (**D**) Spatial distribution of "No change" after an occurrence year.

c. Hot spots and spatial relevance of patterns

One change pattern can be clustered in one region to form a hot spot. In this case, the areas just outside the boundary of the clustered region may display a closely relevant change pattern, and a spatially gradual variation of change patterns emerges in some regions. The two different change patterns share some common form and exhibit spatial relevance. For example, in the Southeast and Central regions, winter maximum temperatures abruptly dropped in the 1970s and then gradually increased afterwards (except in Florida, Illinois and Missouri) (Figure 2.5). However, in Indiana and Ohio, which are right beside the Southeast region, as well as Florida at the boundary of the region, although winter maximum temperatures abruptly dropped in the 1970s, they did not increase significantly afterwards (Figure 2.5). Thus, the two spatially

connected change patterns with relevant forms, “Zero-slope segment followed by a sharp decrease and then an increasing trend” (Type V.D, Figure 2.2) and “Abrupt downward change” (Type IV.D, Figure 2.2) in winter maximum temperature are forming a hot spot in the Southeast and in some Central regions. Hot spots also exist for other climatic variables, such as the 100-year increasing trend of winter maximum temperature and fall total precipitation as shown in Figure 2.3.

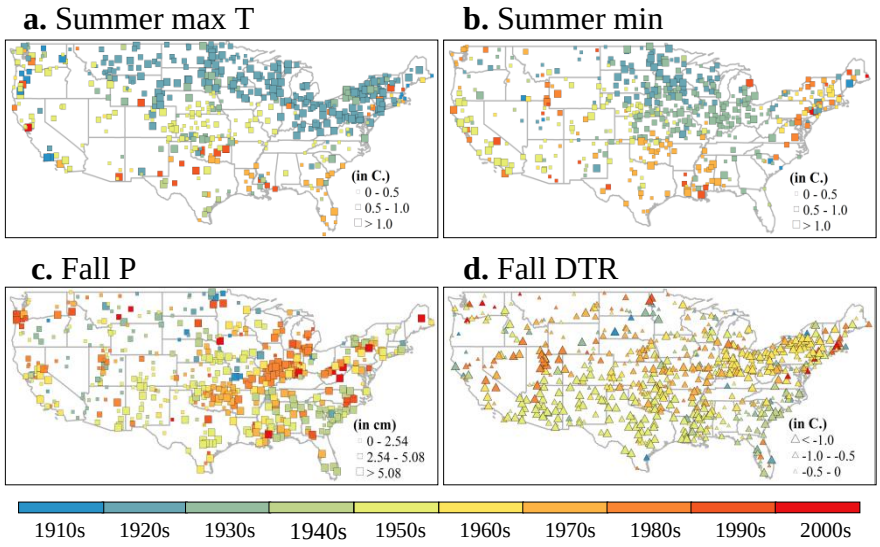


Figure 2.5 Magnitude (shown by sizes of symbols) of an abrupt upward change at occurrence year (shown by colors) for (**a**) Summer mean maximum temperature, (**b**) Summer mean minimum temperature, and (**c**) Fall total precipitation; magnitude of abrupt downward change at occurrence year (shown by colors) for (**d**) Fall mean DTR.

Previous studies presented paradoxical statements about winter temperature change in the Southeast region. For example, studies suggested temperatures did not change in this area from 1950 to 2006 (Easterling et al. 2007); but, at some locations, winter minimum temperature decreased in the Southeast region from 1950 to 2002 (Groisman et al. 2004); and winter mean temperature increased from 1979 to 2005 (IPCC 2013). This paradox may be due to the fact that the abrupt drop in the 1970s was offset by the following gradual increase with the chosen time range ending in 2006. The gradual increase ending in 2002 in some areas may have been

insufficent to offset the effect of the abrupt drop. These detailed changes could not be detected by pure linear regression or simple statistical change tests. One exceptual study, which is from the United States Global Change Research Program (USGCRP), stated that the Southeast annual average temperature had risen with the greatest increase in winter since 1970 (Karl et al. 2009), which is consistent with our findings about the gradual increase trend of winter maximum and minimum temperature after the 1970s.

d. Ongoing changes implying continuous changes

Historical changes detected by records and classified into gradual and abrupt changes provide implications for possible changes at least in the near future. An abrupt change that occurred in the past may or may not occur in the future, while ongoing trends of a climatic variable will very likely continue in the coming years. Therefore, it is more possible for an ongoing trend continued from past decades to continue in coming years than an abrupt change occurring in the past to occur in the near future. This assumption follows similar rules used by IPCC's global warming predictions in terms of CO2 emission and population, i.e., "the models assume that external factors, such as the solar output, continue to behave pretty much the way they have in the recent past" (Chameides 2008).

Figure 2.6 shows the gradual increasing or decreasing trend in the recent past (displayed by red and blue, respectively, in Figure 2.6). The size of symbols in the figure represents the increasing/decreasing rate (i.e., the gradient of the gradual trend), classified into six quantitative categories. In general, a continuing increase in minimum daily temperature is found with more stations than a continuing increase in maximum daily temperature. The findings in our study are consistent with those of Karl et al. (2009): the minimum daily temperature is typically increasing faster or decreasing slower, resulting in a decreasing DTR in most regions in the continental U.S.

There are more cases of increasing trends of temperature than decreasing trends in all four seasons, with an obvious exception for decreasing maximum temperature in the Midwest during summer (Figure 2.6a). In addition, stations without changes dominate in some areas, for example, the summer minimum temperature (Figure 2.6b) and winter/spring maximum temperature in the Midwest (Figure 2.6a). Some findings are consistent with those of previous studies, for example, in part of the Midwest, mean summer temperatures decreased with time during the last century (1895-2010) (Andresen et al. 2012) or the period of 1976-2000 (Pan et al. 2004). The maximum/minimum daily temperature during the winter season is increasing in most stations at the largest rate (Figure 2.6a and 2.6b). In particular, the Southeast region (with the exception of Florida) and the southeast of the Central region are likely to experience an increase of winter maximum/minimum temperature of more than 0.05°C/year. An increasing trend with mean minimum daily temperatures in spring and summer will dominate all regions except the Southeast and the Midwest (Figure 2.6b). During the fall, minimum daily temperature is far more likely to continue to increase than maximum daily temperature, with the largest growth rate in the Central region and the Upper Midwest (Figure 2.6a and 2.6b).

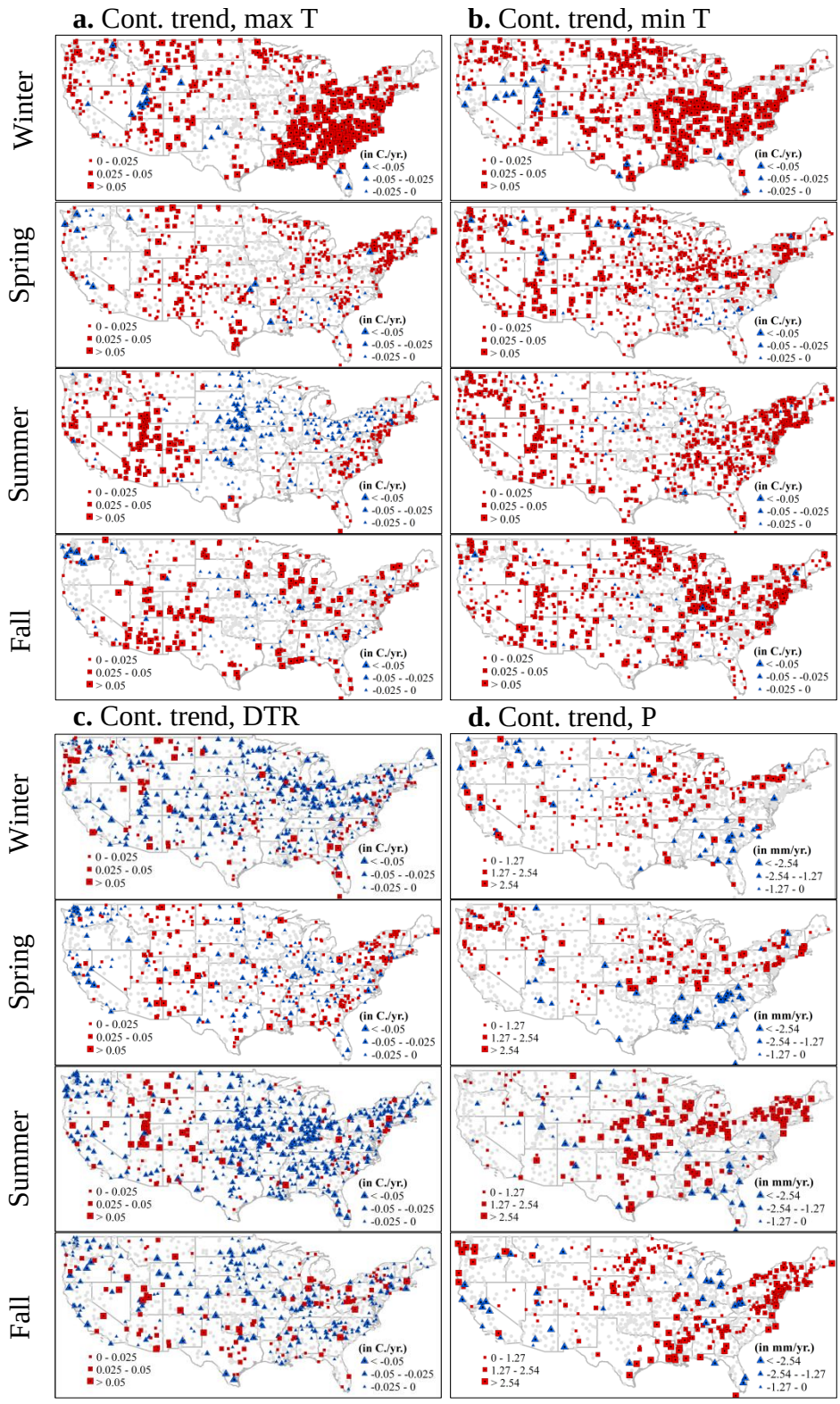


Figure 2.6 Possible continuing increasing (red)/decreasing (blue) trend for the coming years for: (**a**) Seasonal mean maximum temperature, (**b**) Seasonal mean minimum daily temperature, (**c**) Seasonal total precipitation, and (**d**) Seasonal mean DTR in four seasons.

Changes in diurnal temperature range (DTR) vary by season. Winter and summer DTR will likely decrease in more regions than spring and fall DTR. The decreasing gradient of DTR in summer is the largest, especially in the eastern part of the country (Figure 2.6c). The winter DTR in the Upper Midwest and the Central U.S. has been decreasing fast. Based on the identified ongoing trends, spring DTR will possibly continue decreasing in most of the Northwest, West, and Central regions, while it will more likely be increasing in a relatively larger area of the continent, spreading especially in the Northeast, Southeast, and the Mountain regions (Figure 2.6c).

Compared to temperature, few stations show increasing seasonal total precipitation. In winter, spring, and summer, seasonal total precipitation in the Midwest can be projected to increase faster than in other regions, while in the fall, precipitation along the coast of the Atlantic Ocean would likely increase at a high rate (Figure 2.6d). It should be noted that predictions based on climate models are usually made for the long-term future instead of near-term future. Thus, our *speculations* regarding seasonal precipitation change in the coming years can be different from the predictions targeted for the long-term future. For example, some projections suggest that summer precipitation in the Northeast region will have no change or decrease (Hayhoe et al. 2007).

2.4 Conclusions

The identified spatial-seasonal variation and correlation of changes in climate highlights the uncertainty and information loss of averaging climate variables over years and the assumption of linear trends. During 1910-2009, maximum and minimum temperatures in the winter and spring exhibited a much larger possibility of an increasing trend than in the summer and fall. The likelihood of the gradual increasing trend is much larger than that of the decreasing

trend for both maximum and minimum temperature; the abrupt upward change pattern occurred in more locations than the abrupt downward change pattern. However, for DTR, the identified changes are opposite. Abrupt changes in precipitation occurred more frequently than in temperature, and this may continue in the future based on historical occurrences. These abrupt changes probably explain why precipitation change prediction is more uncertain than temperature prediction. Moreover, abrupt changes come with more uncertainty than gradual climate change (Alley et al. 2003), thus the existence of abrupt changes leads to greater difficulties in generating climate scenarios and making projections for impact assessments. With higher unpredictability, abrupt climate change may cause more serious damage on human, ecology, and agriculture systems (Alley et al. 2002).

The ongoing trends identified from historical records imply some continuous changes in the coming years. These ongoing changes are worthy of attention for climate change mitigation and water and land management adaptations now and in the coming years. Compared to the possibly continuing trends, abrupt shifts are less unpredictable probably due to the trigger mechanism, for instance, the Northern American heat wave and drought in 2012 was caused by internal variability of the atmosphere having limited long-lead predictability (Kumar et al. 2013). The unpredictability brings risks in water resources management and agriculture planning.

Moreover, considerable changes in microclimate with heterogeneous change types (see Figure 2.2) and spatial patterns are found in this study, i.e., in nearby areas, climatic variables may experience entirely different change patterns (Figure A.7 in Appendix A). Thus averaging to a regional or global mean will ignore the local heterogeneity and will lose substantial information (Lund et al. 2012). Though global and regional impacts of climate change might still be important, spatial differences in environmental (e.g., urban heat island (Georgescu et al. 2012;

Imhoff et al. 2010)) and societal factors lead the local manifestations and subsequent outcomes (Salih 2012). Thus practices, principles, and policies for the adaption to and mitigation of climate change need to be considered locally, potentially increasing the importance of local governments (Salih 2012) and decision making in local climate change responses.

2.5 References

Alfaro, E. J., Gershunov, A., Cayan, D. (2006), Prediction of summer maximum and minimum temperature over the central and western United States: The role of soil moisture and sea surface temperature, *Journal of Climate 19***,** 1407-1421.

Allard, J., and Keim, B. D. (2007), Spuriously induced temperature trends in the Southeast United States, *Theor. Appl. Climatol. 88***,** 103-110.

Alley, R. B., et al. (2002), *Abrupt Climate Change: Inevitable Surprises.* US National Research Council Report, National Academy Press.

Alley, R. B., et al. (2003), Abrupt climate change. *Science, 299***,** 2005-2010.

Andresen, J., Hilberg, S., Kunkel, K. (2012), Historical Climate and Climate Trends in the Midwestern USA, *U.S. National Climate Assessment Midwest Technical Input Report.*, J. Winkler, J. Andresen, J. Hatfield, D. Bidwell, and D. Brown, Eds.

Bridger, C. A., Ellis, F. P., Taylor, H. L. (1976), Mortality in St. Louis, Missouri, during heat waves in 1936, 1953, 1954, 1955, and 1966, *Environ. Res. 12***,** 38-48.

Chameides, B., Global Warming and Predictions of an Impending Ice Age-Predicting future climate. [Available online at http://blogs.nicholas.duke.edu/thegreengrok/futureclimate/.]

Chernoff, H., and Zacks, S. (1964), Estimating the Current Mean of a Normal Distribution which is Subjected to Changes in Time, *The Annals of Mathematical Statistics 35***,** 999-1018.

Chu, P. S., and Zhao, X. (2011), Bayesian analysis for extreme climatic events: A review, *Atmos. Res*. *102***,** 243-262.

Easterling, D. R., Wallis, T. W. R., Lawrimore, J. H., Heim, R. R. (2007), Effects of temperature and precipitation trends on US drought, *Geophys. Res. Lett*. *34*.

Fong, D. K. H., and DeSarbo, W. S. (2007), A Bayesian methodology for simultaneously detecting and estimating regime change points and variable selection in multiple regression models for marketing research, *Qme-Quant Mark Econ*. *5,* 427-453.

Ge, Y. (2012), Detecting climate change and its impacts on crop yield in the continental United States, Civil & Environmental Eng, University of Illinois at Urbana-Champaign.

Georgescu, M., Mahalov, A., Moustaoui, M. (2012), Seasonal hydroclimatic impacts of Sun Corridor expansion, *Environmental Research Letters*, *7***,** 034026.

Groisman, P. Y., Knight, R. W., Karl, T. R., Easterling, D. R., Sun, B. M., Lawrimore, J. H. (2004), Contemporary changes of the hydrological cycle over the contiguous United States: Trends derived from in situ observations, *J. Hydrometeorol*. *5*, 64-85.

Hamlet, A. F. (2011), Assessing water resources adaptive capacity to climate change impacts in the Pacific Northwest Region of North America, *Hydrol. Earth. Syst. Sc. 15*, 1427-1443.

Hannart, A., and Naveau, P. (2009), Bayesian multiple change points and segmentation: Application to homogenization of climatic series, *Water Resources Research*, *45.*

Hayhoe, K., et al. (2007), Past and future changes in climate and hydrological indicators in the US Northeast, *Clim Dynam*, *28,* 381-407.

Imhoff, M. L., Zhang, P., Wolfe, R. E., Bounoua, L. (2010), Remote sensing of the urban heat island effect across biomes in the continental USA, *Remote Sensing of Environment*, *114*, 504-513.

IPCC (2013) *in Climate Change: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press.

Karl, T. R., Melillo, J. M., Peterson, T. C. (2009), *Global Climate Change Impacts in the United States.* Cambridge University Press.

Kendall, M. G. (1975), *Rank Correlation Methods.* 2 ed. Griffin.

Kim, J., Cheon, S. (2010), Bayesian multiple change-point estimation with annealing stochastic approximation Monte Carlo, *Computation Stat*, *25*, 215-239.

Kumar, A., Chen, M., Hoerling, M., Eischeid, J. (2013), Do extreme climate events require extreme forcings? *Geophys. Res. Lett. 40,* 3440-3445.

Lai, T. L., Xing, H. P. (2011), A Simple Bayesian Approach to Multiple Change-Points. *Stat Sinica*, *21*, 539-569.

Latta, G., Temesgen, H., Adams, D., Barrett, T. (2010), Analysis of potential impacts of climate change on forests of the United States Pacific Northwest, *Forest Ecology and Management, 259,* 720-729.

Lin, J.-G., Chen, J., Li, Y. (2011), Bayesian Analysis of Student t Linear Regression with Unknown Change-Point and Application to Stock Data Analysis, *Computational Economics***,** 1-15.

Lins, H. F., and Slack, J. R. (2005), Seasonal and regional characteristics of US streamflow trends in the United States from 1940 to 1999, *Phys Geogr*, *26*, 489-501.

Lobell, D. B., Schlenker, W., Costa-Roberts, J. (2011), Climate Trends and Global Crop Production Since 1980, *Science, 333,* 616-620.

Lund, M. T., Berntsen, T., Fuglestvedt, J. S., Ponater, M., Shine, K. P. (2012), How much information is lost by using global-mean climate metrics? an example using the transport sector, *Climatic Change*, *113,* 949-963.

Menne, M. J., Williams, C. N., Vose, R. S. (2012), United States Historical Climatology Network Daily Temperature, Precipitation, and Snow Data. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory.

O'Hara, J. K., and Georgakakos, K. P. (2008) Quantifying the Urban Water Supply Impacts of Climate Change, *Water Resources Management*, *22***,** 1477-1497.

Pan, Z. T., Arritt, R. W., Takle, E. S., Gutowski, W. J. Anderson, C. J., Segal, M. (2004), Altered hydrologic feedback in a warming climate introduces a "warming hole", *Geophys Res Lett*, *31.*

Pettitt, A. N. (1979), A non-parametric approach to the change point problem, *Appl. Statist.*, *28*, 126-135.

Ray, B. K., and Tsay, R. S. (2002), Bayesian methods for change-point detection in long-range dependent processes, *J Time Ser Anal*, *23,* 687-705.

Rigaill, G., Lebarbier, E., Robin, S. (2012), Exact posterior distributions and model selection criteria for multiple change-point detection problems, *Stat Comput*, *22,* 917-929.

Rougé, C., Ge, Y., Cai, X. (2013), Detecting gradual and abrupt changes in hydrological records. *Advances in Water Resources*, *53,* 33-44.

Salih, M. A. M. (2012) *Local Climate Change and Society.* Routledge.

Schutz, N., and Holschneider, M. (2011) Detection of trend changes in time series using Bayesian inference, *Phys Rev E, 84.*

Smith, A. F. M. (1975), A Bayesian Approach to Inference about a Change-Point in a Sequence of Random Variables, *Biometrika, 62*, 407-416.

Tai, Y. C., Kvale, M. N., Witte, J. S. (2010), Segmentation and Estimation for SNP Microarrays: A Bayesian Multiple Change-Point Approach, *Biometrics, 66*, 675-683.

Villarini, G., and Smith, J. A. (2010), Flood peak distributions for the eastern United States, *Water Resources Research, 46.*

Villarini, G., Serinaldi, F., Smith, J. A., Krajewski, W. F. (2009), On the stationarity of annual flood peaks in the continental United States during the 20th century, *Water Resources Research, 45.*

Chapter 3. Has irrigation helped or hurt maize production under climate change in Nebraska?

3.1 Introduction

The impact of climate change on crop yields has been a growing concern as a result of global warming and the changes in precipitation patterns (Lobell et al. 2011), and its implications on food insecurity (Thornton et al. 2014). This chapter will use the results of indented changes in climate form Chapter 2 to assess the impact of climate change on the yield of both rainfed and irrigated crops. Nebraska is taken as an example given that the state has substantial irrigated and rainfed crop areas.

Irrigation, which has traditionally been used to deal with water stress and precipitation variability, is expected to mitigate some negative impacts of climate change such as more frequent dryness and extremely high temperatures. The distinction between the impacts on rainfed and irrigated crops is critical since temperature and precipitation affect rainfed and irrigated crops differently, as shown in a study by Troy et al. (2015). Following their study, our study differentiates the impacts of climate change on irrigated and rainfed maize by considering various climate variables, their various change forms (i.e., trends and abrupt changes) and the different crop growth stages. We use maize in Nebraska as an example, using long-term crop yield datasets (1947-2010, which is only available for Nebraska according to the authors' knowledge). We use a longer time series (1910-2010) for the climate dataset, which is needed to detect the change in climate. We develop statistical regression models, which relate crop yields with total precipitation, mean temperature (daily maximum, average and minimum temperature), and mean diurnal temperature range (DTR) for different crop growth stages and for irrigated and rainfed maize separately. We use the regression models to understand how trends and abrupt

shifts in these climate variables in different stages of crop growth affected maize yields during the study period.

We use statistical methods to address the nonlinear relationship of crop yield and climate variables with validation based on historical climate and yield records. While previous assessments have usually assumed that changes in climate occur as gradual trends (Lobell et al 2011), abrupt climatic changes frequently occur, raising important research questions regarding how sudden changes in climate impact human and natural systems (NRC 2013; Rouge et al. 2013). This study examines the impact of both linear and abrupt temperature and precipitation changes on crop yields. Moreover, most of the previous studies have either not differentiated between the impact of climate change on different crop growth stages or have only focused on the crop development stage (Butler and Huybers 2013; Butler and Huybers, 2015; Edreira and Otegui, 2012; Lobell et al. 2013; Ortiz-Bobea and Just, 2012; Sanchez et al., 2014; Schlenker and Roberts 2009). However, these approaches neglect the difference in the sensitivities of crops to climate variables during different stages of crop growth, such as germination and emergence, anthesis and kernel-filling, and ripening (Hu and Buyanovsky 2003). Our method accounts for the impacts of climate change during different crop growing months separately. Understanding the impacts of climate change on different crop stages can help in improving crop models and can be utilized for crop scheduling.

3.2 Materials and Methods

Our regression model is based on a quadratic relationship between the climate variables (i.e., temperature and precipitation) and crop yield, which is an extension of the model of Lobell et al (2011). Crop yield is modeled as a quadratic function of time to reflect technology advancement contribution to crop yield, as well as the impact of climate variables, expressed as:

$$\log(y_{it}) = c_i + b_i t + d_i t^2 + \sum_{k=1}^{5}(\alpha_{1ik} P_{itk} + \alpha_{2ik} P_{itk}^2 + \beta_{1ik} T_{itk} + \beta_{2ik} T_{itk}^2) + \varepsilon_i \quad (12)$$

For county *i* in year *t* (relative to the base year, i.e., 1947 is year *1*), c_i is the intercept reflecting physical conditions (e.g., soil type); b_i and d_i are the coefficients with the time trend; α_{1ik} and α_{2ik} are coefficients for linear and quadratic items of precipitation in five crop growth stages (months) (*k*, *k*=1,..,5 from May to September); P_{itk} is the total precipitation in the k^{th} growth month; β_{1ik} and β_{2ik} are coefficients for linear and quadratic items of a chosen temperature variable and T_{itk} is the mean temperature in the kth growth stage; ε_i is a regression error term.

To examine the impact of different temperature variables, four different regression models are established with the same precipitation variable (monthly total precipitation) but only one of the temperature variables (i.e., T_{max}, T_{min}, T_{ave} and *DTR*). We use the models to quantify the impact of temperature and precipitation changes as follows: first, we perform piecewise fitting of each climate variable based on its original time series from 1910 to 2010 to detect linear trends and abrupt shifts (Ge and Cai, 2018). We then remove (i.e., de-change), the detected changes, in terms of the mean and standard deviation, using the 1947 climate as a baseline to generate the de-changed data (see details of the "de-changing" process in Appendix B). Using the de-changed temperature and precipitation, we then generate three climate scenarios to estimate the crop yield using the above crop yield regression models with: 1) de-changed temperature and actual precipitation, 2) actual temperature and de-changed precipitation and 3) de-changed temperature and precipitation. By comparing the estimated crop yield from each of these climate scenarios with the baseline (i.e., with actual temperature and precipitation), we estimate the crop yield changes due to the change in temperature or precipitation and both temperature and precipitation.

County-level annual irrigated and rainfed maize yield and harvested area, and state-level crop progress dates in Nebraska were obtained from the National Agricultural Statistics Service of the United States Department of Agriculture (USDA). Only those counties (87 out of 93) which have nearly complete maize yield and harvested area data from 1947 to 2010 were chosen for the assessment. The United States Historical Climate Network (USHCN) provides a long-term high-quality data set of daily and monthly observations of precipitation, maximum, average and minimum temperature (Menne et al. 2009). Based on the maize progress dates over the study period in Nebraska, we find that usually 50% of the maize cultivated area is planted by the beginning of May and by the end of September, about 50-75% of the crop is mature. The stages have varied by one to two weeks over the past decades. *Regarding the robustness of the results, we tested different ranges of the growth period.*

Thus, we choose the period of May to September as the growing period in this study. This period is chosen for this study since it results in more significant downward parabolic relationships. We also tested the sensitivity of parameters to the ranges of time. In most counties, the parameter coefficients do not change much for the two periods 1947-2010 and 1970-2010 (see some plots of coefficients in the regression models for irrigated and rainfed maize in Figure B.2 – B.8 in Appendix B). The assumption of normality of model residuals and the issue of autocorrelation of the residuals are addressed in Appendix B.

3.3 Results

3.3.1 Model interpretation

First of all, our model validation shows that the regression models for irrigated maize yield perform considerably better than those for the rainfed yields (about 0.2 higher in the adjusted R^2). This reflects the fact that irrigation stabilizes water availability for maize and

reduces maize yield variability, which results in a smoother relationship between maize yield and the regression variables (climate and technology) (Figure 3.1). Detailed boxplots of each of the parameter coefficients in the regression models are provided in Figure B.2 – B.8 in Appendix B.

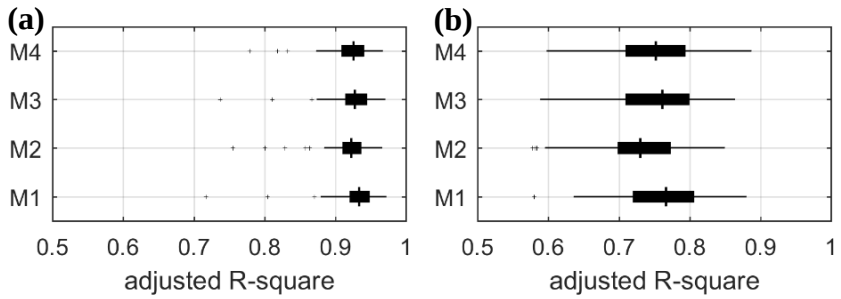


Figure 3.1 Adjusted R-squared for four regression models (M1: P+Tmax, M2: P+Tmin, M3: P+Tavg, M4: P+DTR) in boxplot form: (a) irrigated and (b) rainfed

Consistent with the previous studies (Butler and Huybers 2013; Lobell et al. 2013; Lobell et al. 2011), we find a downward parabolic relationship between both irrigated and rainfed crop yields and climate variables in most counties during the crop growth stages (Table 3.1 for M1 and Table B.1 for M2-M4 in Appendix B), especially in June and July). The downward parabolic relation means that the crop yield first increases with increase in temperature/precipitation but then starts decreasing after a threshold value is reached. Such a nonlinear relationship between temperature and crop yield has also been found by previous studies for other major crops, e.g., soybean (Lobell et al. 2013; Schlenker and Roberts 2009).

We present the thresholds for irrigated and rainfed maize in Table 3.2. During July, more than 70% (80%) of the study counties demonstrate the downward parabolic relationship for *precipitation* and more than 60% (62%) of the counties for the various *temperature* variables for irrigated (rainfed) maize (Table 3.1). The threshold point of the downward parabolic relationship is the temperature or the precipitation, at which the peak value of the corresponding downward parabolic curve is located. The median values of the threshold points over all counties with a

downward parabolic relation are shown in Table 3.2 for the various climate variables (also see details in Figure 3.2). The dominant downward parabolic relationship between the yield and precipitation, especially in July, indicates the negative effect of excessive precipitation (although it might help on water stress to some extent) (Dietzel et al. 2016). The downward parabolic relationship between yield and daily maximum/minimum temperature shows the negative effect of excess heat stress during the day time (Lobell et al. 2013) and high nighttime temperatures (García et al. 2015). As can be seen, the threshold points are different for irrigated and rainfed maize and also vary with the growth stages. For example, the threshold points of mean maximum daily temperature are 24.3 ºC, 28.3 ºC, 28.2 ºC and 28.6 ºC for May, June, July and August for irrigated maize, respectively; while these threshold points are 21.5 ºC, 27.9 ºC, 27.6 ºC and 25.2 ºC for the rainfed maize. As expected, the threshold values for daily maximum temperature are higher for the irrigated crops in the crop growing months as a result of the cooling effect of irrigation, which reduces heat stress of irrigated crops (Lobell et al. 2008). Similarly, the threshold value for precipitation is substantially higher for the rainfed crops than the irrigated crops since the irrigated crops are less dependent on rainfall to meet their water requirements.

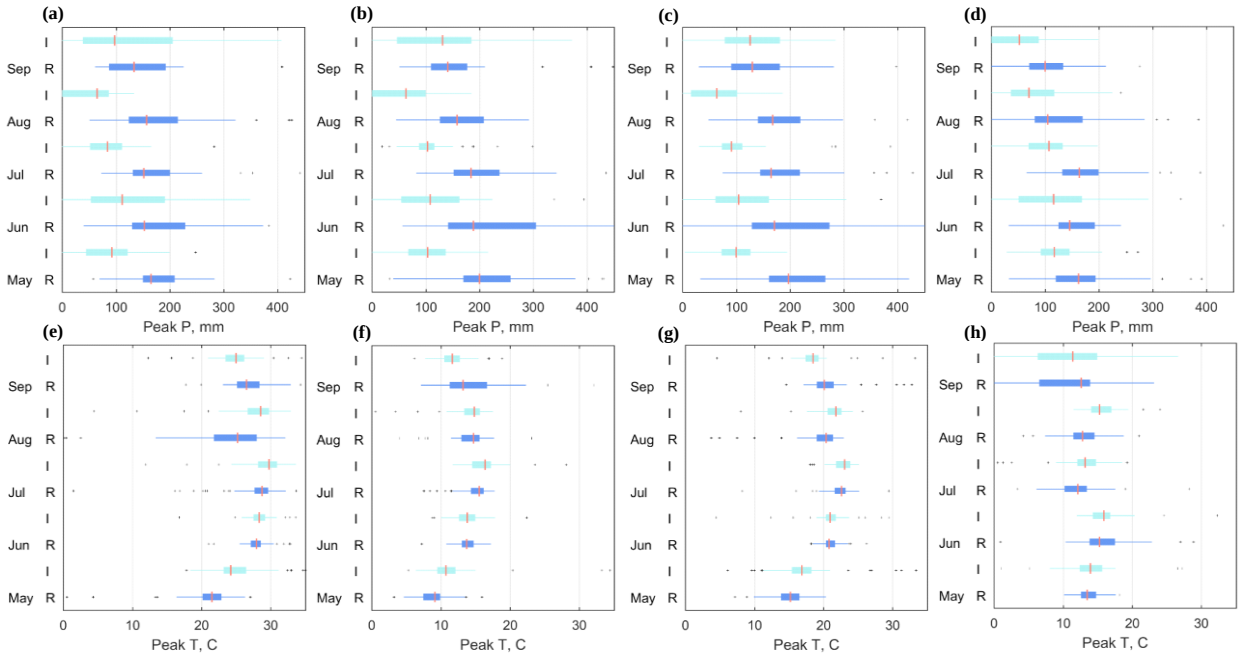


Figure 3.2 The threshold points of precipitation and temperature. Threshold precipitation level and threshold temperature level corresponding to maximum yield, i.e., peak of downward parabolic in five months (May, Jun, Jul, Aug and Sep) in the four regression models ((a)(e) M1: P+Tmax, (b)(f) M2: P+Tmin, (c)(g) M3: P+Tavg, (d)(h) M4: P+DTR) for irrigated (I) maize yield and rainfed (R) maize yield.

Table 3.1 Percentages of counties over all study counties exhibiting the corresponding relationships (upward parabolic and downward parabolic) between crop yield and precipitation or temperature in regresson model M1: P+Tmax

Irrigated (%)	**Relationship**	**M1: P+Tmax**				
Months (k=5, 6, 7, 8, 9)		**May**	**Jun**	**Jul**	**Aug**	**Sep**
Precipitation	Up-parabolic, $\alpha_{2ik} > 0$	49.4	54.0	27.6	57.5	58.6
	Down-parabolic, $\alpha_{2ik} < 0$	50.6	46.0	72.4	42.5	41.4
Temperature	Up-parabolic, $\beta_{2ik} > 0$	12.6	12.6	3.4	23.0	26.4
	Down-parabolic, $\beta_{2ik} < 0$	87.4	87.4	96.6	77.0	73.6
Rainfed (%)	**Relationship**	**M1: P+Tmax**				
Months (k=5, 6, 7, 8, 9)		**May**	**Jun**	**Jul**	**Aug**	**Sep**
Precipitation	Up-parabolic, $\alpha_{2ik} > 0$	48.8	50.0	17.4	29.1	66.3
	Down-parabolic, $\alpha_{2ik} < 0$	51.2	50.0	82.6	70.9	33.7
Temperature	Up-parabolic, $\beta_{2ik} > 0$	25.6	5.8	7.0	48.8	41.9
	Down-parabolic, $\beta_{2ik} < 0$	74.4	94.2	93.0	51.2	58.1

Table 3.2 Median threshold points of each climate variable for irrigated and rainfed maize yield in counties with a downward parabolic relation

Median threshold point	P (mm)*		max T		min T		avg T		DTR	
	Irri	**Rain**	**Irri**	**Rain**	**Irri**	**Rain**	**Irri**	**Rain**	**Irri**	**Rain**
May	103.0	180.8	24.2	21.5	10.7	9.1	16.9	15.2	13.9	13.5
Jun	111.7	164.2	28.3	27.9	13.8	13.7	20.9	20.8	15.9	15.2
Jul	96.1	157.6	28.2	27.6	16.4	15.5	23.1	22.6	13.2	12.1
Aug	64.8	146.7	28.6	25.2	14.8	14.7	21.8	20.4	15.2	12.8
Sep	101.3	125.3	25.0	25.1	11.6	13.2	18.5	20.2	11.4	12.6

* Average of median threshold points for precipitation in the developing monghts over the models with the various temperaure variables.

Using a model that considers precipitation and DTR, we also find a major downward parabolic relationship between the crop yield and DTR for both irrigated and rainfed crops in most of the growth stages. The median value of DTR threshold points of the downward parabolic relationships calculated across all counties are 15.9 ^{0}C and 15.2 ^{0}C in June, 13.2 ^{0}C and 12.1 ^{0}C in July, and 15.2 ^{0}C and 12.8 ^{0}C in August for irrigated and rainfed maize, respectively. Crop yields respond best under a certain level of DTR (Draper 1998; Le 2011) with a positive (negative) relationship between crop yield and DTR below (above) the threshold point. Irrigated maize has a higher threshold DTR than rainfed maize since irrigation effectively reduces the range of daily maximum and minimum temperature experienced by the crops (Bonfils and Lobell 2007; Kanamaru and Kanamitsu 2008, Kueppers et al. 2008).

3.3.2 Impact assessment of climate change on crop yield

Based on the established regression models and generated climate scenarios, we assess the impact of changes in monthly temperature and precipitation within the growing period. We

try to understand how the changes in crop yields observed in the historical data are associated with the abrupt shifts and/or gradual trends in mean and/or variance detected in the climate variables (see how abrupt shifts and gradual trends and variance are detected in Chapter 2).

We selected two counties from Nebraska to demonstrate the characteristics of changes in climate variables and yields, which were observed in most parts of the state. Figure 3.3 is used to discuss the characteristics of changes in June precipitation observed in Nebraska, by showing the results for Polk County in eastern Nebraska. Figure 3.4 shows the results for Thayer County in southeastern Nebraska, which is used to discuss the changes in July precipitation observed in most counties, while Figure 3.5, also for Polk County, is used to demonstrate the changes in maximum, mean and minimum temperature and DTR in the state. Furthermore, Table 3 (associated with Figures 3.3-3.5) displays the cumulative percentage change of crop yield (weighted by annual harvested area) in both Polk County and Thayer County during the study period, as a result of changes in each of the climate variables.

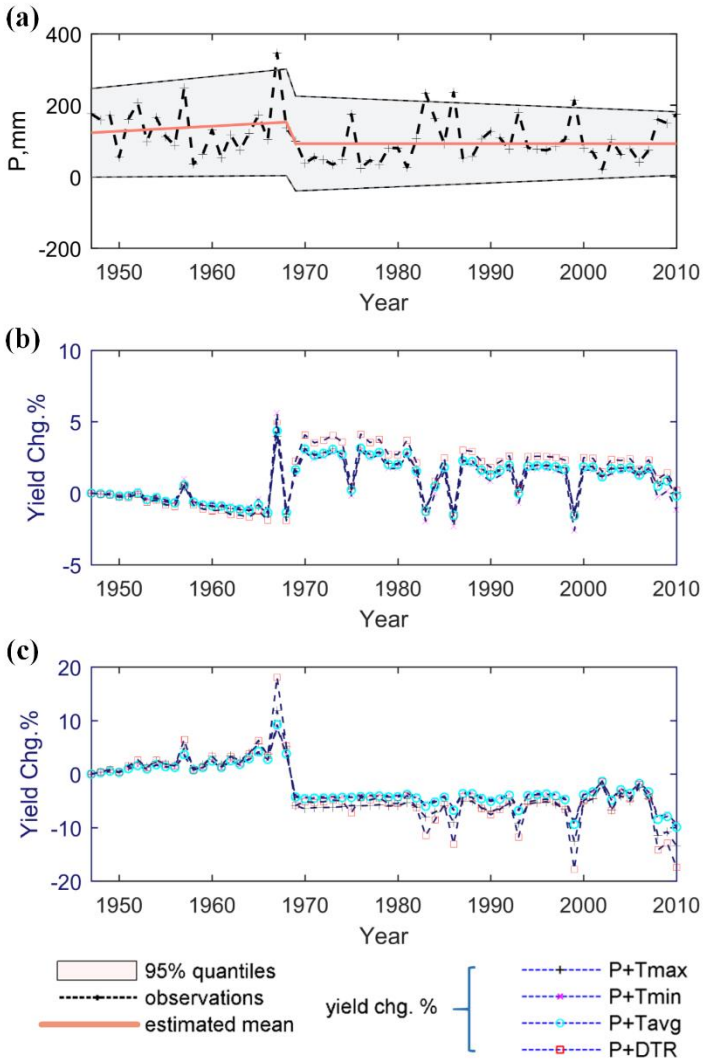


Figure 3.3 Precipitation change over time and corresponding annual changes in irrigated and rainfed maize yield in Polk County. (a) The changes in June precipitation. (b) The irrigated maize yield change due to precipitation change. (c) The rainfed maize yield change due to precipitation change. Precipitation and one type of temperature are regressed with log yield in four models (i.e., P+Tmax, P+Tmin, P+Tavg and P+DTR)

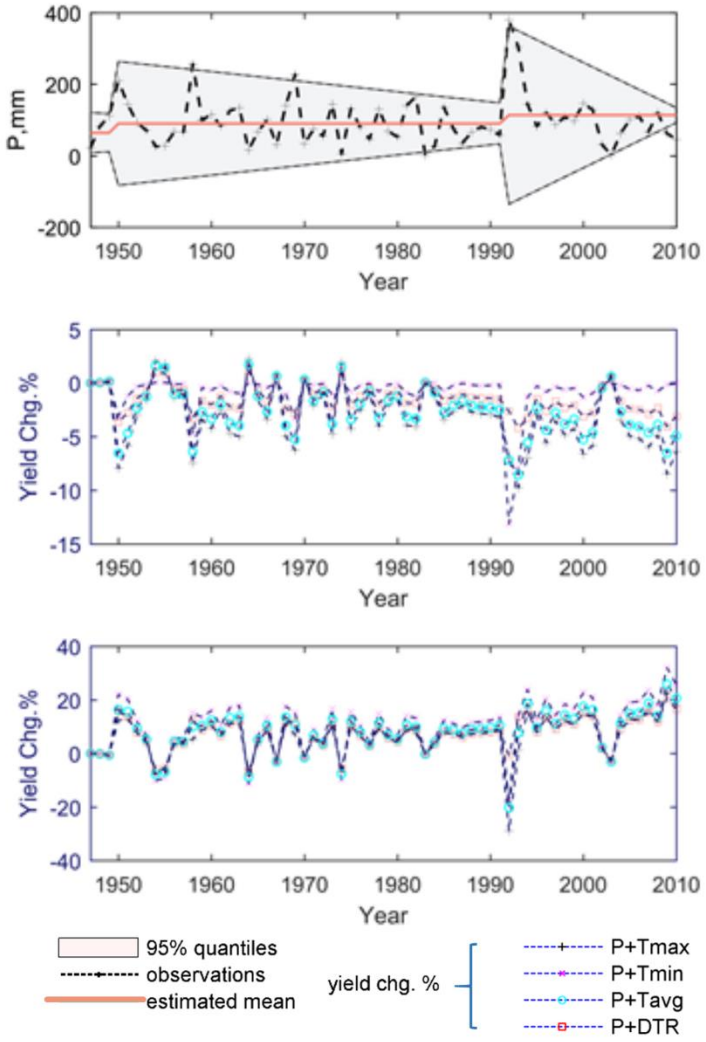
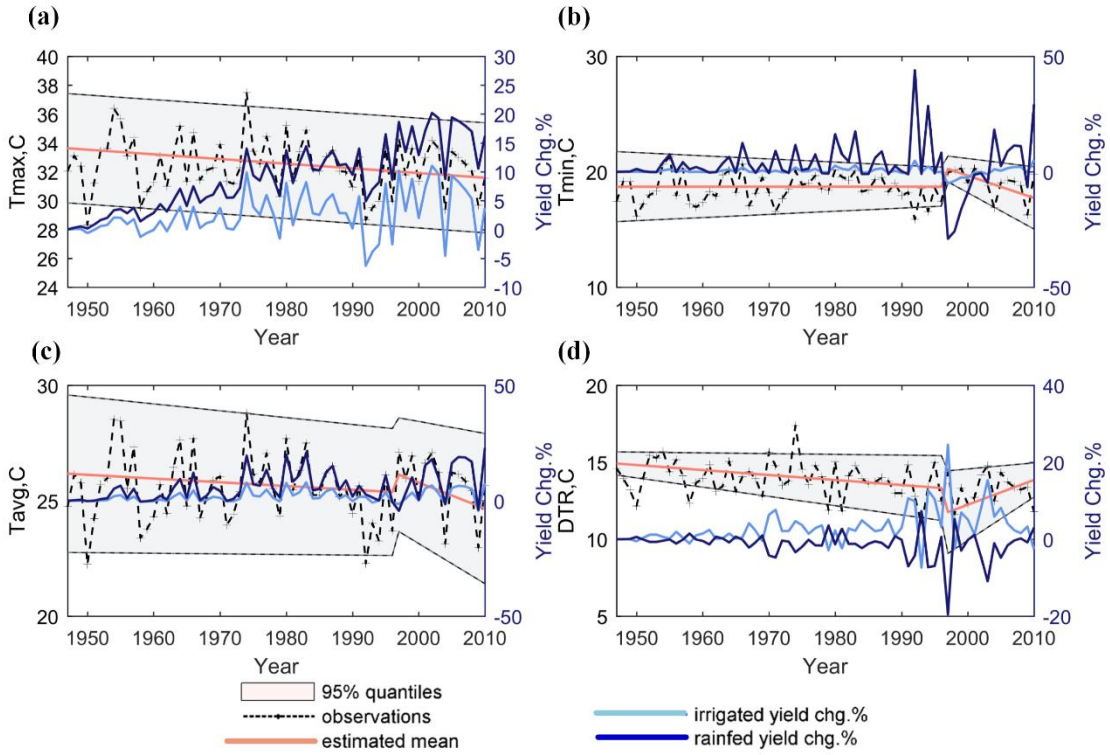


Figure 3.4 Precipitation change over time and corresponding annual changes in irrigated and rainfed maize yield in Thayer County. (a) The changes in July precipitation. (b) The irrigated maize yield change due to precipitation change. (c) The rainfed maize yield change due to precipitation change. Precipitation and one type of temperature are regressed with log yield in four models (i.e., P+Tmax, P+Tmin, P+Tavg and P+DTR)



* 95% quantiles and estimated mean of DTR are those of daily maximum temperature minus the corresponding values of daily minimum temperature.

Figure 3.5 Temperature variable change over time and corresponding annual changes in irrigated and rainfed maize yield in Polk County. The changes in July mean temperature variables (a. *Tmax*, b. *Tmin*, c. *Tavg* and d**.** *DTR*) and how they impacted irrigated and rainfed maize yield. Precipitation and one type of temperature are regressed with log yield with each of the four models

Table 3.3 Cumulative contribution of temperature and precipitation change in June or July over the period 1947-2010 to irrigated and rainfed maize yield change (%) in Polk County and Thayer County, Nebraska (associated with Figs. 3-5). They are in four regression model between log yield and precipitation and one type of temperature

Cumulative irrigated yield chg.% due to	**P+Tmax**	**P+Tmin**	**P+Tavg**	**P+DTR**
Jun Precipitation chg.	1.4	0.02	1.4	1.9
Jul Precipitation chg.	-4.0	-1.0	3.2	-1.9
Jul Temperature chg.	3.3	0.34	2.3	4.4
Cumulative rainfed yield chg.% due to				
Jun Precipitation chg.	-2.2	-3.5	-2.8	-3.9
Jul Precipitation chg.	10.5	16.1	12.5	10.7
Jul Temperature chg.	9.4	2.0	4.9	-1.5

In the month of June, an increasing trend in precipitation was observed during 1947-1967, which resulted in an increase in the rainfed maize yield but a decrease in the irrigated yield (Figure 3.3). This is due to the fact that the threshold point of the downward parabolic relationship for precipitation is lower for irrigated maize (Table 3.2). Additional June precipitation over the threshold point for irrigated maize adversely affects irrigated maize yield, although irrigated maize yields still remain higher than the rainfed yields due to water stress mitigation by irrigation and other agricultural inputs going along with irrigated crops. This implies that while on one hand, the soil moisture maintained in the irrigated crop fields is beneficial for crop growth; on the other hand, the high soil moisture during the irrigation season, can aggravate excessive soil wetness after strong storms, and cause crop loss due to root rots and

Pythium and Bacterial Stalk Rot (Kanwar 1988). Studies have found increase in the number of extreme climate events, which involve alternating excessive wetness and dryness, especially the situation of dryness quickly followed by wetness in the Midwestern United States (e.g., Fishman 2016; Hatfield et al. 2011; Rosenzweig et al. 2002). The dry events can encourage famers to irrigate the farms, but if the dry event is followed by high rainfall, it can lead to excessive wetness of soil, which can damage the crops.

An abrupt decline of June precipitation around 1968, resulted in an abrupt decrease of rainfed maize yield, but an increase of irrigated maize yield (Figure 3.3). The abrupt change in 1968 was again followed by a gradual increase in precipitation in the subsequent years, which again decreased the irrigated yields and increased the rainfed yields. The gross effect of changes in June precipitation over all study years appears to be positive for irrigated maize but negative for rainfed maize in Polk County (Table 3.3) and in most counties of Nebraska. There is an overall positive impact on irrigated maize yields (about 66% of all counties) and a negative impact on rainfed maize yields (about 55% of all counties) (Figure 3.6). This shows the offsetting of the negative effects on irrigated maize caused by gradual changes in precipitation before and after 1968, by the positive effects due to the abrupt change in 1968. However, if we only assess the long-term trend during the study period, neither a significantly positive nor a significantly negative trend can be detected for the precipitation in June, and such analysis shows no impact of precipitation change on crop yield (Figure B.9 in Appendix B). This validates the importance of considering the impacts of abrupt changes in climate on crop yield.

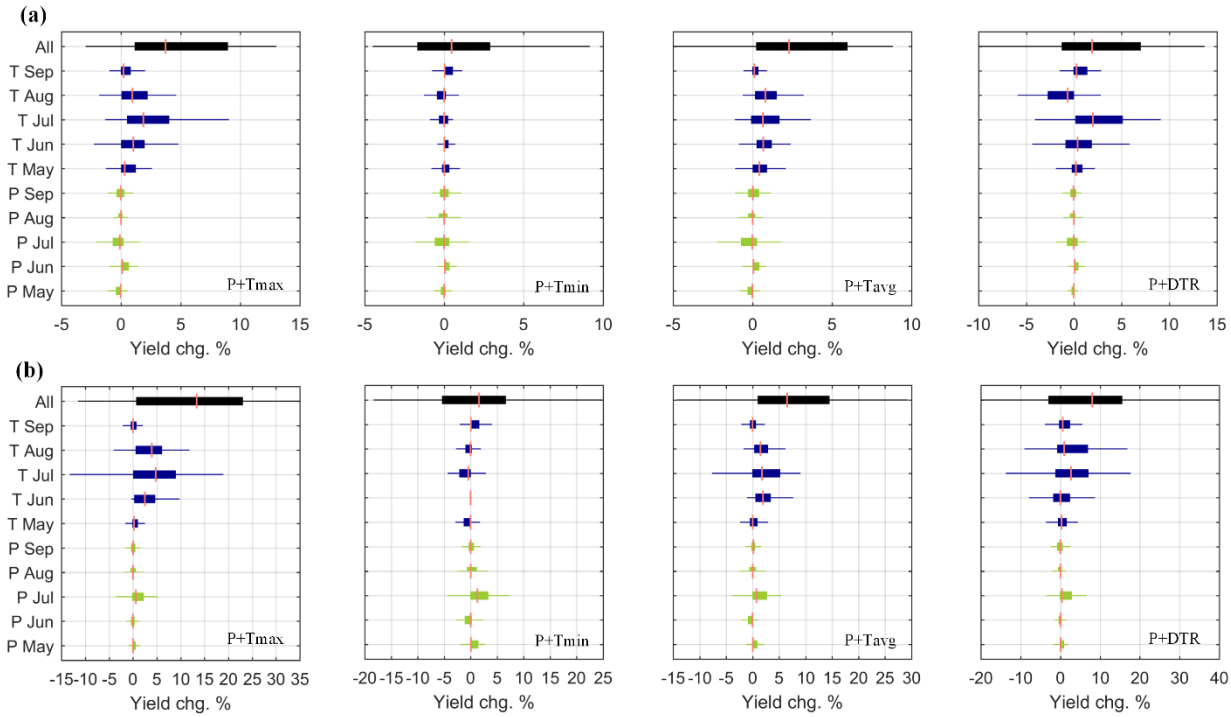


Figure 3.**6** Cumulative percentages of yield change due to precipitation and temperature change in all counties over the whole study period. They are in the five growth months (stages) (May, Jun, Jul, Aug and Sep) in the four models and the overall impact from both temperature and precipitation change for (a) irrigated and (b) rainfed maize over the study counties in Nebraska

In July, the gross effect of the changes in precipitation in most of the counties is opposite to that of changes in June precipitation (Figure 3.6). July precipitation in Thayer County presents the typical change pattern of July precipitation over all counties in Nebraska. It abruptly increased around 1949 and 1991, with largely decreasing variance after each of the abrupt shifts (Figure 3.4a). Similar to the increased June precipitation in Polk County, the increased July precipitation in Thayer County benefited the rainfed yield but hurt the irrigated yield, except in some years with extreme precipitation. For example, the extreme precipitation in 1992, which was over the threshold points of both the irrigated and rainfed maize, had a negative effect on both the types of crops. The low variance after 1995 benefited rainfed maize. Overall, the gross effect of changes in July precipitation in Thayer County was negative for irrigated maize and considerably positive to rainfed yield (Table 3.3). Similar changes in July precipitation occurred

in most of the counties, though some had only one abrupt increase around 1991 and some experienced a linear increase over the entire study period. Overall, the median impact of changes in July precipitation over all Nebraskan counties was negative for irrigated maize and positive for rainfed maize (Figure 3.6).

Temperature variables also exhibited abrupt shifts and trends in the observation period. An abrupt increase in daily minimum (Figure 3.5b) and daily average (Figure 3.5c) temperature was observed in July around the year 1995, as a result of which the maize yields, especially the rainfed maize yield, experienced an abrupt decline and did not begin to recover until around 2005. After that, maize yields showed an increasing trend associated with the declining trend of the minimum and average temperature. A significant decreasing trend in maximum temperature was observed in the study period (Figure 3.5a), which clearly had a positive impact on both rainfed and irrigated crop yields. It is worth noting that the rate of increase in yield of rainfed maize was larger than that for irrigated maize in most of the years, i.e., 2% and 0.34% for rainfed and irrigated maize, respectively, accounting the accumulated effects over the entire study period (Table 3.3), although the annual variability of rainfed maize yield was larger than that of irrigated maize.

The July DTR in Polk County (Figure 3.5d and Table 3.3) decreased from a level below the threshold point for rainfed yield but above the threshold point for irrigated yield (note that the threshold point for rainfed crops is higher than that for irrigated crops in this county, which represents an exception to the median value in Table 3.2). As a result, the decrease in DTR was accompanied by a decrease in rainfed maize yield but an increase in irrigated maize yield. Even though the changes in DTR in most counties were similar to that observed in Polk County (linearly decreasing trend followed with an afterward abrupt downward shift or a linearly

decreasing trend), its impact on maize yields were different from the case of Polk County. In those counties, the July DTR decreased from a level higher than the threshold level for both rainfed and irrigated maize. The decrease in DTR led to an increase in both the rainfed and irrigated maize yields, but with higher increase for rainfed crops because of lower DTR threshold point for rainfed crops (e.g., Banner County in Figure B.11 in Appendix B). The median increase in crop yields due to the changes in July DTR was 2.62% for rainfed maize as compared to 1.99% for irrigated maize (Figure 3.6).

Overall, most counties in Nebraska showed an increase in maize yields throughout the study period. The impacts on maize yields were larger from the changes in temperature (*Tmax* and *DTR*) than from the changes in precipitation (Figure 3.6). The changes in maximum temperature in July had the most significant impact on the changes in crop yields among the climate variables considered. The decreasing trend in July mean daily maximum temperature resulted in an increase in the crop yields in most of the counties (as shown in the case of Polk County, Figure 3.5a), except for a few counties in southwest Nebraska (e.g., Chase County and Banner County). In those counties, maximum and/or minimum temperature abruptly increased/decreased around 1991 and linearly increased to values above the threshold points, resulting in gross negative effect for both rainfed and irrigated maize yields (Figure B.10 for Chase County in Appendix B). However, it is possible that the positive effect from the decreased maximum temperature cannot be offset by the negative effect from the increasing maximum temperature as the increasing trend is not long enough (for example, Figure B.11 in Appendix B for irrigated maize in Banner County), which results in gross positive effect.

3.4 Discussion and Conclusions

In this study, we used statistical models to understand the influence of climate variables on crop yields of rainfed and irrigated maize in Nebraska during 1947-2010. The statistical models were used to determine how abrupt and gradual changes in temperature and precipitation affected the yields of rainfed and irrigated crops in Nebraska in different stages of crop growth of the observation period. We found significant differences in the relationship between crop yields and climate variables for rainfed and irrigated crops. We also found that the impacts of changes in climate variables on crop yields vary significantly with the stages of crop growth. Our study adds to the knowledge from previous studies, that evaluated the entire crop growth season or only summer season (e.g., Butler and Huybers 2013; Lobell et al. 2013; Schlenker and Roberts 2009), by making critical distinctions between different crop growth stages and by distinguishing between irrigated and rainfed maize.

We found that the changes in temperature had a more significant influence on the changes in crop yields in the observation period as compared to the changes in precipitation. The impact of temperature and precipitation changes were more significant in the crop development stage (June and July) as compared to the planting and harvesting stages. Among the various climate variables considered, a strong decreasing trend in daily maximum temperatures in the month of July had the most significant impact on the yields of both rainfed and irrigated crops. This decrease in maximum temperature increased the yields of both the rainfed and irrigated crops, but rainfed crops were benefitted more. For precipitation too, the most significant impact on crop yields was observed in the month of July. The changes in July precipitation benefitted the rainfed yields but contributed towards a decrease in the irrigated yields. In June, the changes in precipitation had a positive influence on the irrigated yields and a negative influence on the

rainfed yields, however, the impacts of changes in June precipitation were much smaller as compared to the impacts of July precipitation. Thus, the precipitation changes during the whole growing period benefitted the rainfed crops more than the irrigated crops. Overall, the results of our study show that rainfed crops were benefitted more from the changes in the climate in the observation period than irrigated crops.

The effectiveness of adaptive management measures such as changes in crop variety, planting date, crop density, changes in the level of fertilization, and application of irrigation (Cuculeanu et al. 1999) can be enhanced if they are implemented in such a way that they can take advantage of the impacts of climate change on crop yields. In this study we found that irrigation was overall beneficial to maize production in Nebraska, but the expected returns were reduced due to the impacts of climate change. Our results show that the benefits of irrigation in Nebraska can be enhanced if it is applied by taking into account the changes in temperature and precipitation that have been taking place in the region. For example, irrigation is known to reduce the adverse impact of excess heat through its cooling effect. However, the fact that there has been a significant long-term decrease in maximum temperatures in July, which has been benefitting rainfed crops more than irrigating crops, shows that the amount of irrigation can be reduced to take advantage of the cooling. Also, abrupt increases in precipitation were observed in the month of July, which had a negative effect on irrigated crop yields due to excessive soil wetness. The frequency of extreme dry events followed by wet events has been increasing in the Midwest region (Fishman 2016; Hatfield et al. 2011; Rosenzweig et al. 2002), which can increase the risk of excessive soil wetness due to irrgiation in the dry periods. Enhanced investment in integrated irrigation and drainage systems and in seasonal forecasting of precipitation might be ways to reduce such losses.

Furthermore, the results from this study provide knowledge for developing new or updating existing relationships in climate-crop models especially for the purpose of assessing the impact of climate change, including 1) the downward parabolic relationship between crop yield (both rainfed and irrigated) and mean monthly climate variables, 2) the various threshold points of maximum, average and minimum temperature and precipitation for crop growing months, and 3) the different impacts of those variables on irrigated and rainfed maize in the study area.

3.5 References

Bonfils, C., Lobell, D. (2007), Empirical evidence for a recent slowdown in irrigation-induced cooling, *Proc Natl Acad Sci 104*, 13582-13587.

Butler, E. E., Huybers, P. (2013), Adaptation of US maize to temperature variations, *Nature Clim Chang 3*, 68-72.

Butler, E. E., Huybers, P. (2015), Variations in the sensitivity of US maize yield to extreme temperatures by region and growth phase, *Environ Res Lett 10*, 034009.

Cuculeanu, V., Marica, A., Simota, C. (1999), Climate change impact on agricultural crops and adaptation options in Romania, *Clim Res 12*, 153-160.

Dietzel, R., Liebman, M., Ewing,. R et al (2016), How efficiently do corn- and soybean-based cropping systems use water? A systems modeling analysis, *Glob Chang Biol.* doi:10.1111/gcb.13101

Draper, N.R. (1998), Applied regression analysis bibliography update 1994-97, *Commun Stat - Theor M 27*, 2581-2623.

Edreira, J. I. R, Otegui, M. E. (2012), Heat stress in temperate and tropical maize hybrids: Differences in crop growth, biomass partitioning and reserves use, *Field Crop Res 130*, 87-98.

Fishman, R. (2016), More uneven distributions overturn benefits of higher precipitation for crop yields, *Environ Res Lett 11*, 024004

García, G. A., Dreccer, M. F., Miralles, D. J. et al (2015), High night temperatures during grain number determination reduce wheat and barley grain yield: a field study, *Glob Chang Biol 21*, 4153-4164.

Ge, Y., Cai, X. (2018), Abrupt shifts and gradual trends of temperature and precipitation of the past century in the continental United States, under revision.

Hatfield, J. L., Boote, K. J., Kimball, B. A. et al. (2011), Climate Impacts on Agriculture: Implications for Crop Production, *Agron J 103*, 351-370.

Hu, Q., Buyanovsky, G. (2003), Climate effects on corn yield in Missouri, *J Appl Meteor 42*, 1626-1635.

Kanamaru, H., Kanamitsu, M. (2008), Model Diagnosis of Nighttime Minimum Temperature Warming during Summer due to Irrigation in the California Central Valley, *Journal Hydrometeor 9*, 1061-1072

Kanwar, R. S., Baker, J. L., Mukhtar, S. (1988), Excessive soil water effects at various stages of development on the growth and yield of corn, *Trans Am Soc Agric Engineers 31*, 133-141

Kueppers, L. M., Snyder, M. A., Sloan, L. C., et al. (2008), Seasonal temperature responses to land-use change in the western United States, *Glob Planet Chang 60*, 250-264.

Le, P. V. (2011), Climate change impact on crop yields: Reexamine evidence from decreasing diurnal temperature range, European Association of Environmental and Resource Economists 18th Annual Conference.

Lobell, D. B., Schlenker, W., Costa-Roberts, J. (2011), Climate Trends and Global Crop Production Since 1980, *Science 333*, 616-620.

Lobell, D.B., Bonfils, C. J., Kueppers, L. M. et al. (2008), Irrigation cooling effect on temperature and heat index extremes, *Geophys Res Lett 35*, L09705.

Lobell, D. B., Hammer, G. L., McLean, G. et al. (2013), The critical role of extreme heat for maize production in the United States, *Nature Clim Chang 3*, 497-501.

Menne, M. J., Williams, C. N., Vose, R. S. (2009), The Us Historical Climatology Network Monthly Temperature Data, Version 2, *B Am Meteorol Soc 90*, 993-1007.

NRC (National Research Council) (2013), Abrupt impacts of climate change: Anticipating surprises. Committee on Understanding and Monitoring Abrupt Climate Change and its Impacts (eds White JW, Alley RB, Archer DE et al), the National Academies Press, Washington DC

Ortiz-Bobea, A., Just, R. E. (2012), Modeling the structure of adaptation in climate change impact assessment, *Am J Agr Econ.* doi: 10.1093/ajae/aas035

Rosenzweig, C., Tubiello, F. N., Goldberg, R., et al. (2002), Increased crop damage in the US from excess precipitation under climate change, *Glob Environ Chang 12*, 197-202.

Rouge, C., Ge, Y., Cai, X. (2013), Detecting gradual and abrupt changes in hydrological records, *Adv Water Resour 53*, 33-44

Sanchez, B., Rasmussen, A., Porter, J.R., (2014), Temperatures and the growth and development of maize and rice: a review, *Glob Change Biol 20*, 408-417

Schlenker, W., Roberts, M. J. (2009), Nonlinear temperature effects indicate severe damages to US crop yields under climate change, *Proc Natl Acad Sci 106*, 15594-15598

Thornton, P.K., Ericksen, P.J., Herrero, M., et al. (2014), Climate variability and vulnerability to climate change: a review, *Glob Chang Biol 20*, 3313-3328

Troy, T. J., Kipgen, C., Pal, I. (2105), The impact of climate extremes and irrigation on US crop yields, *Environ Res Lett 10*, 054013.

Chapter 4. Drought pattern change underlying climate change

4.1 Introduction

Drought is a recurrent extreme climate phenomenon. It can last for weeks, months, even years, and the spatial extent of droughts is usually larger than other natural hazards (e.g., floods and hurricanes) (Obasi, 1994), resulting in devastating impacts on agriculture, water resources, environment and human lives (Hao et al., 2014; WMO, 2006). Droughts differ significantly in terms of spatial characteristics from one region to another. Drought identification and quantification are prerequisites to drought frequency analysis. Numerous drought indices are used to quantify drought events, including the widely used PDSI (Palmer Drought Severity Index) and SPI (Standardized Precipitation Index) (Mishra and Singh, 2010), and recent new indices, e.g., SPEI (Standardized Precipitation Evapotranspiration Index) (Masud et al., 2015), and SDI (streamflow drought index) (Madadgar and Moradkhani, 2013; Sadri and Burn, 2014). PDSI is probably the most popular regional drought index to monitor droughts and to assess agricultural impacts (Mishra and Singh, 2010). Recent advances in technologies use improved methods, e.g., using the Penman-Monteith equation instead of Thornthwaite to calculate Potential Evapotranspiration (PET) (Dai, 2011); using self-calculating technologies (Wells et al, 2004), etc.); accounting for snowmelt using degree-day model (van der Schrier et al., 2013); more accurate climate datasets have also been used to improve the quantification of PDSI at the global scale (Sheffield et al., 2012). Those updated PDSI data sets have been used to explore how drought is changing under climate change. However, there have been conflicting results from recent studies (Dai, 2013; Sheffield et al., 2012; Trenberth et al., 2014; van der Schrier et al., 2013) due to different forcing climate datasets, methods involved in calculating potential evapotranspiration, self-calculating periods, and other factors.

Drought occurs event by event; different events last a wide range of time periods and some can last over multiple years. Therefore, it is not appropriate to assess drought return periods using the classic procedures applied to intra-year extreme events (e.g., floods, Shiau and Shen, 2001), which only consider annual maximum. The length of a drought event and the cumulative severity can be more important than the maximum intensity of the event for assessing its impact on natural and human systems (e.g., agriculture and water demand). Thus, it has been suggested to use multiple characteristics including expected drought inter-arrival time, severity (S), duration (D), and peak intensity (I) for drought assessment (Shiau, 2006; Shiau and Shen, 2001). Given that the multiple drought characteristics are correlated, recent studies propose to assess drought return periods based on those correlated variables. Copulas (Sklar, 1959) based on joint multivariate probability distributions allow modeling a multivariate distribution by separately dealing with marginal distributions and joint dependences among variables (Madadgar and Moradkhani, 2013; Sadri and Burn, 2014; Wong et al., 2010; Xu et al., 2015). Thus it has been widely used to analyze return periods of the various extreme hydroclimatic variables including peak flow and water volume for rainfall frequency (Zhang and Singh, 2007), flooding events (De Michele et al., 2005) and drought events. Specifically, bivariate copulas has been applied to calculating drought return period based on the combination of two factors among D, S and I using the time series of the various indices, such as SPI (Chen et al., 2013; Lee et al., 2013; Wong et al., 2010), SPEI (Masud et al., 2015), and SDI (Madadgar and Moradkhani, 2013; Sadri and Burn, 2014). Moreover, integrated drought indices based on multivariate copulas are suggested, e.g., probability-based overall water deficit index from multiple drought-related indices (Kao and Govindaraju, 2010), multivariate drought index utilizing information from multiple hydroclimatic variables (Rajsekhar et al., 2015) and integrated multivariate standardized

drought index (e.g., standardized Palmer drought index-based joint drought index, SPDI-JDI) (Ma et al., 2014; Ma et al., 2015). Especially, in India, copulas have been applied to deriving drought severity-duration-frequency or intensity-area-frequency curves based on SPI using bivariate copulas in two western states (i.e., Gujarat, western Rajasthan) (Ganguli and Reddy, 2012; Reddy and Ganguli, 2012; Reddy and Ganguli, 2013). Mishra and Singh (2010) reviewed mostly used drought indices and concluded that PDSI based on an inherent time scale is suitable for assessing agricultural impacts (e.g., Quiring and Papakyriakou, 2003; Lee and Nadolnyak, 2012; Yan et al., 2016). Drought characteristics (e.g., frequency and severity) based on monthly PDSI were used to generate a drought risk index to assess the relationship between drought and crop yield reduction (Li et al., 2009). Understanding the change in multiple drought characteristics (D, S and I) is needed to examine both the separate and joint effect of the characteristics on agriculture.

The IPCC AR5 concludes that there is not enough evidence available in favor of or against any global trend in drought with high confidence, and admits that the global increasing trend in drought suggested by the IPCC AR4 was probably overestimated (IPCC, 2013). The same concern presents itself with the various assessments at the regional scale.

In addition, meteorological drought events vary significantly from one region to another due to different climate characteristics. Internal variability in the Earth's climate causes the temporal variation of droughts, which has also been linked to climate change (Burke et al., 2006; Dai, 2013; Cook et al., 2015). Understanding the temporal and spatial characteristics of droughts can help in evaluating future drought risk and in choosing appropriate drought mitigation strategies. The spatial patterns of drought characteristics are analyzed at a range of time scales from month to year (Vicente-Serrano, 2006). In addition, annual time series of drought characteristics have

also been analyzed using streamflow records to detect spatial and temporal changes (Clausen and Pearson, 1995; Hisdal et al., 2001). However, few of these studies has considered the dynamic changes of drought characteristics including both spatial and temporal variability. Although drought characteristics (D, S and I) are correlated, the spatial and temporal patterns of these characteristics may not be the same (Shiau and Shen, 2001; Shiau, 2006). Trend analysis applied to drought characteristics D, S and I can determine whether drought events become longer, more severe, or more intense. This analysis can reveal more information than a trend analysis that is applied directly to a drought index time series (e.g., Dai 2013; Ficklin et al 2015) and may even lead to different results. An increasing or decreasing trend in a drought index (e.g., Palmer Drought Severity Index (PDSI)) does not necessarily mean an increasing or decreasing trend in D, S and I. For example, a region with decreasing D may have an increasing trend in S (Andreadis and Lettenmaier, 2006). More recent studies (Sheffield et al., 2009; Ganguli and Ganguly, 2015; Liu et al., 2016) studied multiple drought characteristics (D, S and I), but none of these studies have analyzed the trends for each drought characteristic individually.

Analysis of trends of D, S or I should be associated with a particular return level (i.e., the frequency of a certain level of drought events). This is because drought events with different return periods may have different spatial and temporal patterns. For example, if we consider two locations, one may have a higher severity under a 5-year return period while the other may have drought events that have a higher severity under a 50-year return period. Thus, in this study, we analyze trends in drought characteristics under a given return period, covering a series of return periods from shorter to longer ones. Some recent studies have also analyzed drought characteristics separately for different return periods (e.g., Ganguli and Ganguly, 2015; Liu et al., 2016).

There are two studies in the following sections accessing drought characteristics in two regions, e.g., India and U.S.

4.2 Drought Frequency Change: An Assessment in Northern India Plains

In India, drought occurs mostly due to the failure of south-west monsoon (June-September) (Reddy and Ganguli, 2012). About 33% of the arable land in India is considered to be drought-prone and a further 35% can also be affected under extreme climate conditions (Ganguli and Reddy, 2012). Northern and eastern India is a major agriculture area of India, especially for wheat and rice production. Drought frequency analysis in this region is needed to conduct risk evaluation and select drought-relief measures.

The purpose of this study is to provide a more comprehensive drought assessment using PDSI at the regional level (i.e., northern and eastern India). Compared to previous studies (Dai, 2013; Sheffield et al., 2012) that focus on trend analysis in time series of a drought index, our study distinguishes the changes in multiple drought characteristics. We use a time series of drought events to examine 1) the characteristics of the time series such as D, S and I of drought events, and 2) the changes associated with both individual and joint drought characteristics. Using the copulas to simulate the joint distribution of multiple drought variables, we will assess both the univariate and multivariate return periods of drought during three historical periods (1900-1954 and 1955-2012 of Dai (2013) and 1948-2008 of Sheffield et al. (2012)). We will conduct the assessment in eight areas in our study region located in northern India plains. Due to data availability and the advantage of PDSI that better help assess agriculture impacts (Mishra and Singh, 2010), we choose PDSI as the drought index for the assessment. Based on the outputs, we attempt to understand the change in bivariate return period of both D and S or both D and I over thresholds. Specifically, we will address the following questions: 1) has severe drought

become more frequent, longer-lasting, or more intense over the historical period? 2) How are the two datasets (Dai and Sheffield et al.) distinguished, in severe or mild droughts, or in long or short droughts? 3) Did the multivariate return period change over time and space?

4.2.1 Data and study region

4.2.1.1 The characteristics of a drought time series

PDSI is a meteorological drought index using precipitation and temperature for estimating moisture supply and demand within a two-layer soil model (Palmer, 1965). The basis of the index is the difference between the amount of precipitation required to retain a normal water-balance level and the amount of actual precipitation (supply-demand concept of water balance). Figure 4.1 shows the sketch of a time series of monthly PDSI. A drought event has four major components (Saghafian and Mehdikhani, 2014): a) duration (D) expressed in months, during which drought index is continuously below a prescribed critical level; b) drought inter-arrival time (T) expressed in months, which is the time range between the initiation time of two consequent drought events; c) severity (S) indicating the cumulative deficiency of a drought event below the critical level; and d) peak intensity (I), which indicates the maximum absolute value of a monthly drought index below the critical level.

Eleven states for drought and moist events are identified with PDSI: extremely wet, very wet, moderately wet, slightly wet, incipient wet spell, near normal, incipient drought, mild drought, moderate drought, severe drought and extreme drought (Palmer, 1965). To focus on relatively severe drought events, the truncation level is set to -1 (corresponding to “mild drought”) in this study, which means that only the events with PDSI less than -1 are collected to form the time series for the drought assessment in this study. Drought characteristics including D, S, I and T are calculated based on monthly PDSI time series for each drought event.

Since a monthly time series is used to represent PDSI, D is discretized by an interval of one month, by which a number of drought events can have the same D value. The presence of events with an identical value of a variable is called “ties”, especially with small D values. The univariate fitting distributions and bivariate fitting distributions (copulas) in this study require continuous random variables. Therefore, the integer numbers of D are converted into continuous decimal numbers by using the randomization scheme proposed by De Michele et al. (2013) and Xu et al. (2015), which adds a random variable that is uniformly distributed in the range of [-0.5, +0.5] to the primary integer values of D. It is proved that this randomization process maintains statistical characteristics of the original data (De Michele et al., 2013).

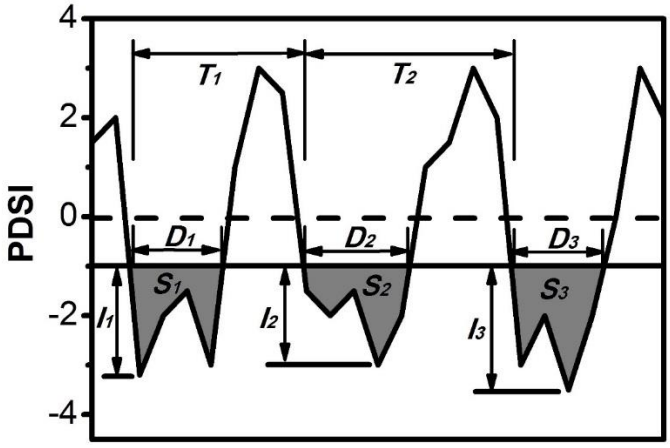


Figure 4.1 Definition sketch of drought events; D_i is the duration, S_i is the severity, $I_i$is the peak intensity and T_i is the inter-arrival time.

4.2.1.2 Data Sources and Case Study

In the study, we use two different PDSI databases: 1) The dataset used by Dai (2013) with 2.5 degree spatial resolution from 1900 to 2012; 2) The dataset used by Sheffield et al. (2012) with 1 degree spatial resolution from 1948 to 2008. Our study region is northern and eastern India (Figure 4.2). Due to the different resolutions of the data sets, we match the grids of

the two sources to extract most comparable data. Time series of monthly PDSI extracted from the dataset of Dai and Sheffield et al. are plotted for the eight chosen grids in Figure 4.3. Comparison of PDSI time series from the two datasets is shown in the figure with considerable differences especially in recent decades in areas 2-5.

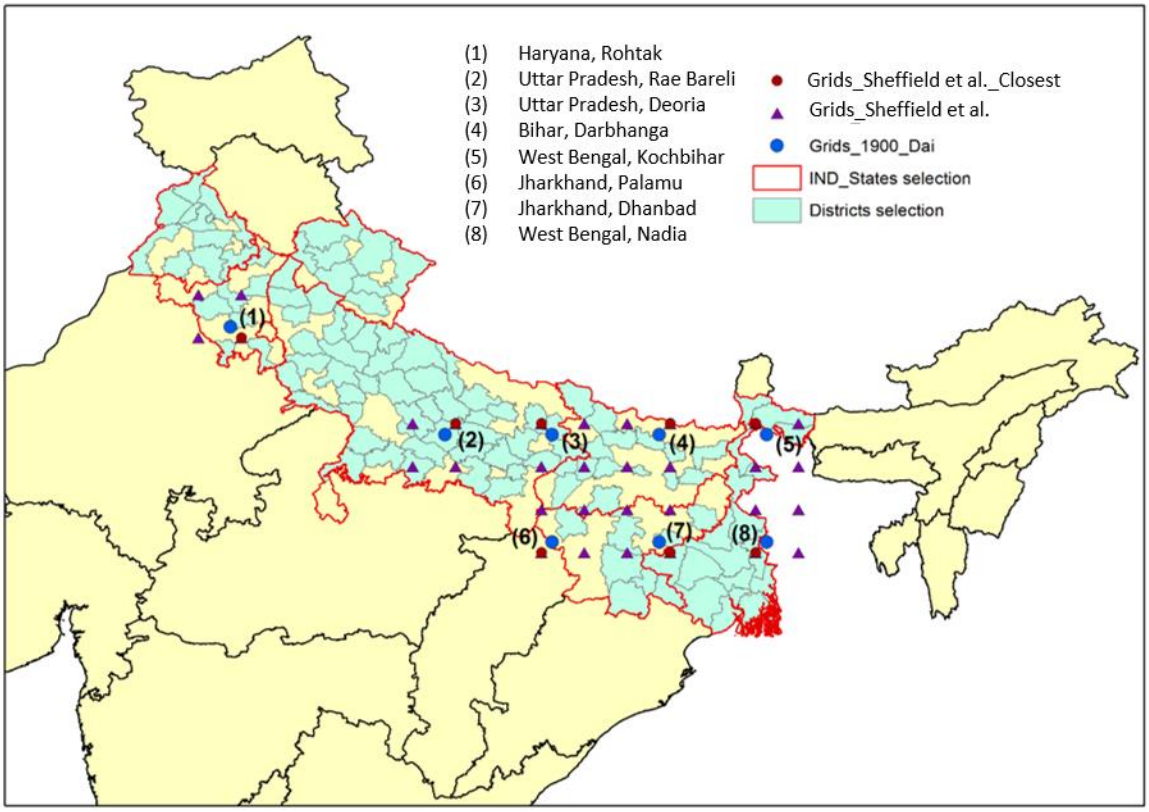


Figure 4.2 Spatial locations of the study areas in the northern and eastern India. Blue dots indicate the eight PDSI grids in Dai's dataset (Dai, 2013), numbered from 1 to 8. Red dots are chosen as the closest grid in Sheffield et al. (2012) to the eight grids in Dai's dataset, respectively.

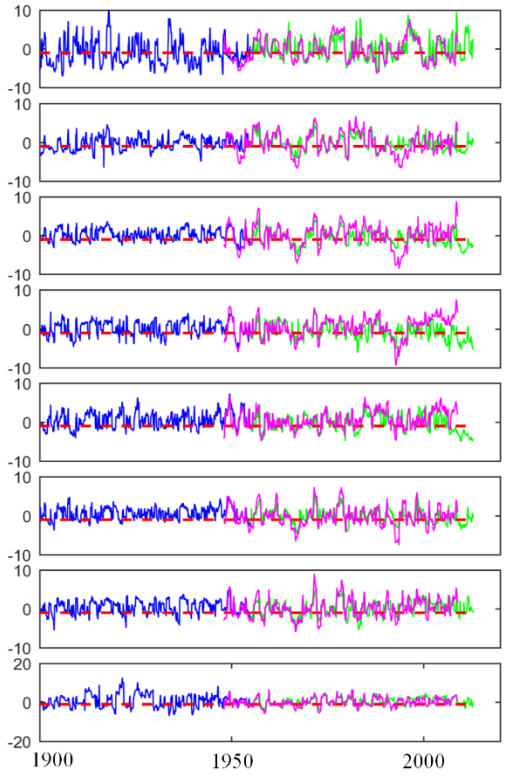


Figure 4.3 Time series of monthly PDSI: eight chosen grids from Dai's dataset from 1900 to 2012 compared to that from Sheffield et al.'s from 1948 to 2008, extracted from a grid mostly matched to the data grid of Dai. Blue: 1900-1954, green: 1955-2012, magenta: 1948-2008.

4.2.2 Methods

4.2.2.1 Fitting of Univariate and bivariate probability distributions

(1) Selected univariate distributions

The 'best' estimated fitted distributions based on KS test at the 99% ($\alpha = 0.01$) significant level are chosen for each area in each period. The fittings of D, S and I are acceptable in all areas in all periods. The randomization procedure dealing with "ties" in D improves the fitting. Meanwhile, based on the agreement between empirical CDF (based on the original data) and theoretical CDF (Figure 4.4a), it is proved that the procedure doesn't affect the statistical characteristics of the original data. Also, the agreement between empirical and theoretical CDFs in all areas in all periods is visually supported as shown in Figure 4.4b (for S) and Figure 4.4c

(for I). The mostly selected fitting distributions for D, S and I are Pearson type III, Weibull and Generalized Pareto, respectively. In previous studies, Exponential (Lee et al., 2013; Shiau, 2006), Log-normal, and Weibull (Madadgar and Moradkhani, 2013; Wong et al., 2010), Gamma distributions (Chen et al., 2013) are commonly used for D, the Gamma distribution (Lee et al., 2013; Madadgar and Moradkhani, 2013; Shiau, 2006) for S, and Gamma (Madadgar and Moradkhani, 2013), Pearson Type III, and the Generalized Pareto distributions (Chen et al., 2013) for I.

4.2.2.2 The joint probability distribution of D, S and I based on copulas

An n-copula is a multivariate cumulative distribution function (CDF) with all the univariate marginal distributions being uniform on the interval [0,1], i.e., C: $[0,1]^d \rightarrow [0,1]$ (Nelsen, 2007). Copula obeys Sklar's theorem (Sklar, 1959), i.e., there exists an n-copula C binding univariate margins $F_1, \dots, F_n$ to a joint distribution function (H):

$$H(x_1, x_2, \dots, x_n) = C[F_1(x_1), F_2(x_2), \dots, F_n(x_n)] = C(\mu_1, \mu_2, \dots, \mu_n) \quad (13)$$

where $F_k(x_k) = \mu_k$, for $k = 1,2, \dots, n$ with $\mu_k \sim U[0,1]$. The copula approach allows modeling a multivariate distribution by separately dealing with marginal distributions and joint dependences among variables (Madadgar and Moradkhani, 2013; Wong et al., 2010; Xu et al., 2015). The theorem suggests separating the development of a multivariate joint distribution model for *H* into two parts: 1) estimation of the marginal CDFs, F_k, and 2) estimation of the copula C.

In this study, we choose eight commonly used univariate probability distributions as the candidate margins to fit the distributions for D, S and I: Exponential, Gamma, Weibull, Log-normal, Generalized Logistic, Generalized Pareto, Generalized Extreme Value and Pearson type III. The estimation of the univariate distribution uses the L-moments method (R package *lmomco*)

(Asquith, 2015). Finally, Kolmogorov-Smirnov (KS) test (Marsaglia et al., 2003) is used to select the best fitted distribution.

The commonly used Elliptical copula (Normal and *t*) and Archimedean (Clayton, Gumbel and Frank) are selected as the candidate copulas. The inversion of Kendall's τ (Nelsen, 2007; Xu et al., 2015) method is used to estimate the parameters of these copulas. The parametric bootstrap based Cramer-von Mises test (Genest et al., 2009) is used for testing the goodness of fit. The bivariate joint distributions are constructed using the R package *copula* (Hofert and Maechler, 2011; Hofert et al., 2015; Kojadinovic and Yan, 2010; Yan, 2007).

4.2.2.3 Return period in a bivariate framework

Derived from the classic return period, the drought return period is associated with a certain exceedance probability. However, a drought event may happen multiple times in one year or last for more than one-year long. Therefore, Shiau and Shen (2001) and Shiau (2006) generated a formula to calculate drought return period with expected drought inter-arrival time (T):

$$T_D = \frac{E(T)}{1-F_D(d)} \quad (14)$$

$$T_S = \frac{E(T)}{1-F_S(s)} \quad (15)$$

$$T_I = \frac{E(T)}{1-F_I(i)} \quad (16)$$

where T_D, T_S, and T_I are return periods with a single variable, D, S, or I, greater than or equal to a certain value, respectively; $F_D(.), F_S(.)$ and $F_I(.)$ are percentiles of CDFs with D, S, and I, respectively; and $E(T)$ is the expected inter-arrival time of sequential droughts within the study period.

Considering the drought return period based on dependent two variables, the joint distribution of those variables takes the form of copulas. The return period for bivariate is defined as the return period for "D and S": $T_{\cap(DS)}$ for D ≥ d and S ≥ s (Shiau, 2003). The joint return period for a drought event with copula-based joint distribution is described as:

$$T_{\cap(DS)} = \frac{E(T)}{P(D \geq d, S \geq s)} = \frac{E(T)}{1 - F_D(d) - F_S(s) + F_{DS}(d,s)} = \frac{E(T)}{1 - F_D(d) - F_S(s) + C[F_D(d), F_S(s)]} \quad (17)$$

The joint bivariate return periods for "D and I" ($T_{\cap(DI)}$) can be formulated in a similar form of Equation (17).

4.2.3 Results

4.2.3.1 Fitting of Univariate and bivariate probability distributions

(2) Selected univariate distributions

The 'best' estimated fitted distributions based on KS test at the 99% ($\alpha = 0.01$) significant level are chosen for each area in each period. The fittings of D, S and I are acceptable in all areas in all periods. The randomization procedure dealing with "ties" in D improves the fitting. Meanwhile, based on the agreement between empirical CDF (based on the original data) and theoretical CDF (Figure 4.4a), it is proved that the procedure doesn't affect the statistical characteristics of the original data. Also, the agreement between empirical and theoretical CDFs in all areas in all periods is visually supported as shown in Figure 4.4b (for S) and Figure 4.4c (for I). The mostly selected fitting distributions for D, S and I are Pearson type III, Weibull and Generalized Pareto, respectively. In previous studies, Exponential (Lee et al., 2013; Shiau, 2006), Log-normal, and Weibull (Madadgar and Moradkhani, 2013; Wong et al., 2010), Gamma distributions (Chen et al., 2013) are commonly used for D, the Gamma distribution (Lee et al., 2013; Madadgar and Moradkhani, 2013; Shiau, 2006) for S, and Gamma (Madadgar and

Moradkhani, 2013), Pearson Type III, and the Generalized Pareto distributions (Chen et al., 2013) for I.

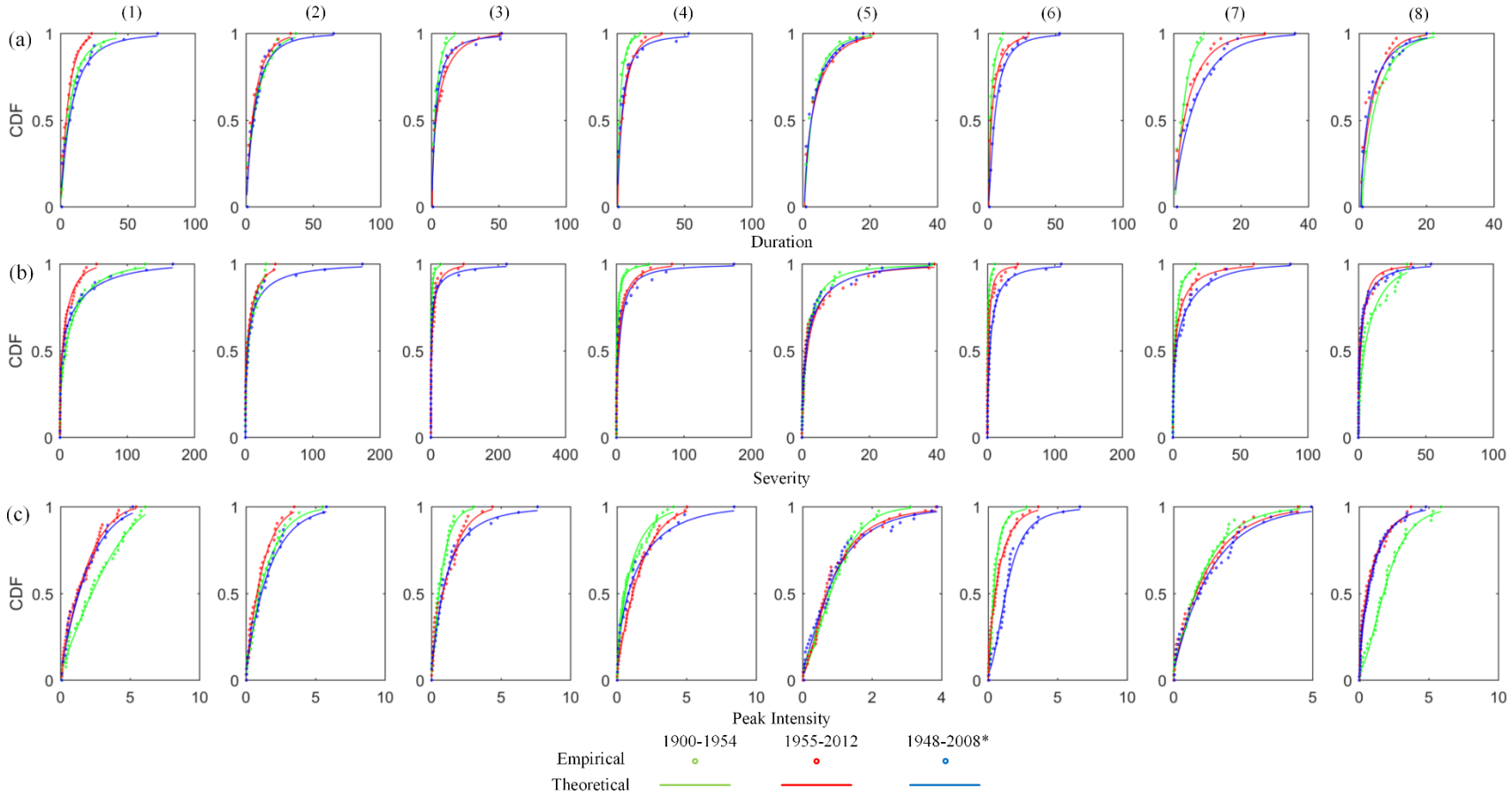


Figure 4.4 Comparison of empirical CDF with the 'best' estimated distribution, i.e., theoretical distribution for PDSI (a) duration (D), (b) severity (S) and (c) intensity (I) in eight areas in northern and eastern India, based on the dataset of Dai (2013) for the periods: 1900-1954 and 1955-2012, and the dataset of Sheffield et al. (2012) for the period of 1948-2008.

(3) Selected bivariate copula functions

Correlation between D and S and between D and I can be observed (not shown here) from the scatter plots for the pair of D-S and D-I for all periods and areas. The correlated D-S and D-I approve the necessity of using copulas-based drought frequency analysis methods. The "ties" are resolved using the randomization method described in section 2.1.

The "best" fitted bivariate copula functions for pairs of D-S and D-I in all eight areas and in three periods are selected by p-value of the CM test. Owing to the randomization procedure dealing with "ties" in D, the fitting of pairs D-S and D-I has been significantly improved. The test shows that the fitted distributions of pairs of D-S and D-I cannot be rejected for any area and any period at the significance level as 0.01. The agreement of the empirical and theoretical CDFs is displayed in Figure 4.5. The most fitted bivariate copula functions are the Frank copula and Clayton copula; while Gumbel copula is chosen in previous bivariate copula study of SPI in northern India (Ganguli and Reddy, 2012; Reddy and Ganguli, 2012; Reddy and Ganguli, 2013).

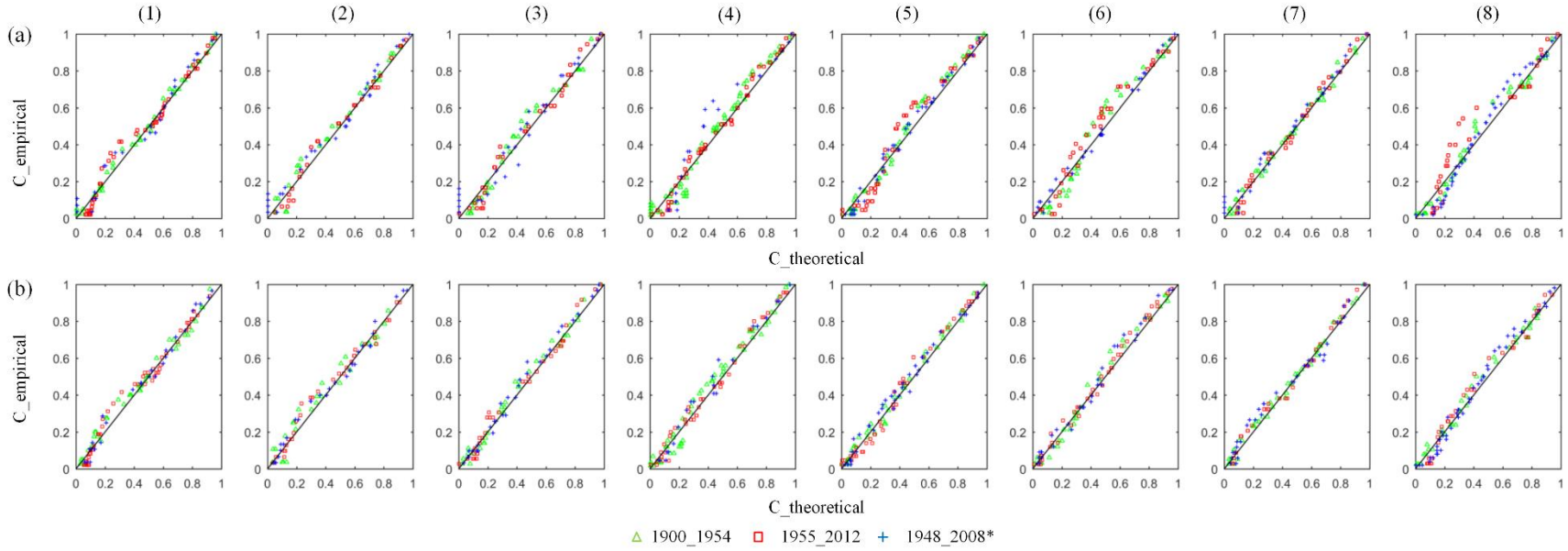


Figure 4.5 Comparison of empirical CDF with the 'best' estimated joint distribution of (a) D-S and (b) D-I, i.e., theoretical copulas for D-S and D-I of PDSI in eight areas in northern and eastern India, based on the dataset of Dai (2013) for two periods: 1900-1954 and 1955-2012, and the dataset of Sheffield et al. (2012) for period 1948-2008.

4.2.3.2 Return periods

(1) Univariate return period

Firstly, temporal and spatial variations of univariate return periods are shown in Figure 4.6 based on the univariate return period using the data set of Dai (2013). Figure 4.6 plots the comparisons of the univariate return periods in terms of D, S and I during two historical periods (1900 to 1954, 1955 to 2012), respectively, for the eight areas in northern and eastern India. The figure shows how the return periods vary by different drought characteristics, over time and space. The D, S and I for a given drought return period was lower after 1955 in areas 1, 2 and 8 (except that the severity stayed almost the same level in area 2), for example, the 100-year drought severity was lower in the later historical period in areas 1, 2 and 8; while D, S and I were all higher after 1955 in areas 3-7 (Table 4.1). The large decrease of D and S under almost all return periods after 1955 in area 1 is probably due to a significant increasing trend in precipitation and a decreasing trend in PET (Madhu et al., 2015). Meanwhile the large increase of D and S in area 3, 4, 6 and 7 is due to the significant decreasing trend in precipitation but relatively stable PET (Madhu et al., 2015).

The mean monthly PDSI values (only those below -1 are chosen for the comparison) are calculated (Table 4.1) for the growing seasons of wheat and Kharif rice (the two major crops in Northern and Eastern India): Nov.-Mar. for wheat and Jun.-Oct. for Kharif rice (AQUASTAT, 2016). It shows consistent changes in D, S and I under the 10-, 50- and 100-year return period in most of the study areas. In areas 3-7, there had increasing D, S and I associated with decreasing (the opposite direction) of PDSI calculated for the growing period of Kharif rice; while in areas 1, 2 and 8, there had decreasing D, S and I (except for 10-year S in area 2) with increasing PDSI calculated for the growing period of wheat. Areas 1-2, located in Northern India, plant 38% of

the crop land for wheat compared to 18% for rice. Areas 3-8, located in Eastern India, plant 46% of their crop land for Kharif rice and 34% for wheat (AQUASTAT, 2016). The situation of longer, more severe and more intense drought in areas 3-7 may threaten the irrigated rice planted in those areas; while the improved drought situation can be beneficial for wheat in areas 1 and 2. Crop pattern change (e.g., reducing rice acreage and increasing wheat acreage in areas 3-7) can be an adaptive measure.

The frequency analysis with D, S and I is partially consistent with trend analysis of precipitation and PDSI conducted by Dai (2013), who found that precipitation increased and PDSI increased (becoming wetter) in areas around area 1 and area 8 in the historical periods. However, the same trend analysis by Sheffield et al. (2012) shows the PDSI in areas around area 1 and area 2 decreased. Such trend analysis in PDSI can reveal how monthly PDSI changes over time, but cannot distinguish the changes in different drought variables (i.e., D, S and I), neither the correlation of those variables with both long-term and short-term drought events, e.g., higher S for long-term drought events but lower S for short-term drought events after 1955 in area 2, as shown in this study.

To further show the changes that vary by areas, Figure 4.7 compares the eight areas in each of the three historical periods, revealing considerable spatial difference with the return periods for each of the three drought variables. Under the same return period, during 1900-1954, area 1 had the longest D, the highest S and I, followed by areas 2 and 8, while area 6 had the lowest or second lowest level for all the three drought variables. Duration and severity in area 7 were relatively low, but peak intensity was at a medium level over all the areas in period 1900-1954. In the following period (1955-2012), under the same return period, area 3 had the longest

D and highest S; area 4 had a most severe drought with all the three aspects; area 6 was larger in D but was still small in S and I compared to period 1900-1954.

Moreover, the comparisons between the assessments based on period 1955-2012 of Dai (2013) and period 1948-2008 of Sheffield et al. (2012) are also shown in Figure 4.6 for D, S and I. The duration is longer and the severity is higher in Sheffield et al (2012) than Dai (2013) under almost all return period years, especially in longer-term return period years, except area 5. The differences of peak intensity level under all return period years are smaller in all areas, but a bit larger in Sheffield et al (2012) than Dai (2013) especially in longer-term return period years. As the severity is the cumulative intensity during a drought event, the difference between two data sets is due to the duration of the identified drought events.

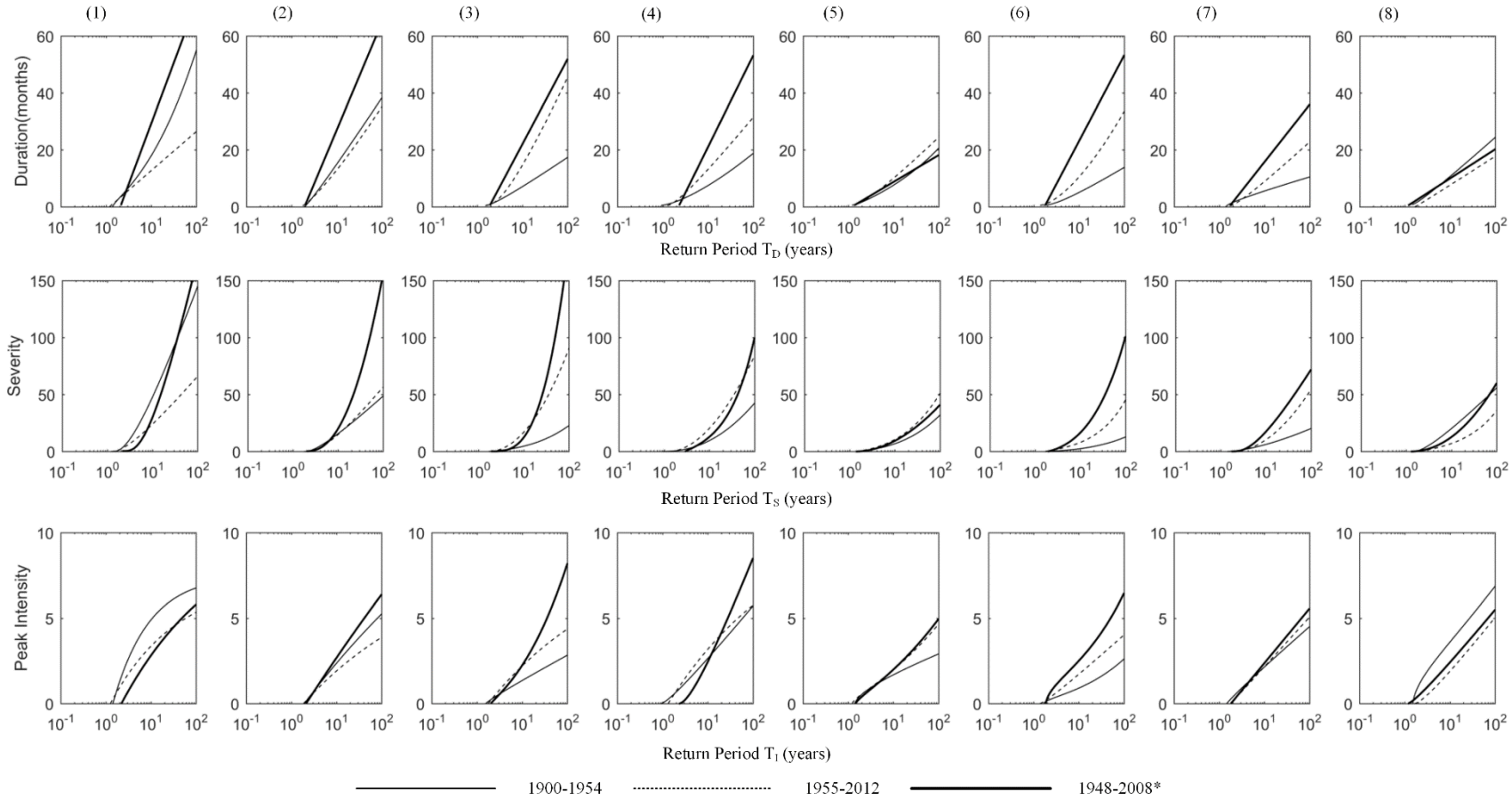


Figure 4.6 Return period in terms of D, S and I for dataset of Dai (2013) during two historical periods (1900 to 1954, 1955 to 2012), respectively, and for the dataset of Sheffield et al. (2012) during the period of 1948 to 2008 for the eight areas in northern and eastern India.

Table 4.1 Comparison of D, S and I levels under 10-year, 50-year and 100-year return periods and comparison of mean monthly PDSI in growing seasons of wheat and Kharif rice during the two historical periods using Dai (2013)

Area #	1	2	3	4	5	6	7	8
D	-	-	+	+	+	+	+	-
S	-	+*	+	+	+	+	+	-
I	-	-	+	+	+	+	+	-
Mean monthly PDSI								
Wheat, Nov-Mar	0.81	0.23	-0.06	-0.01	0.03	-0.12	0.00	0.05
Kharif Rice, Jun-Oct	0.68	0.19	-0.24	-0.20	-0.17	-0.20	-0.17	0.50

* - for the 10-year return period

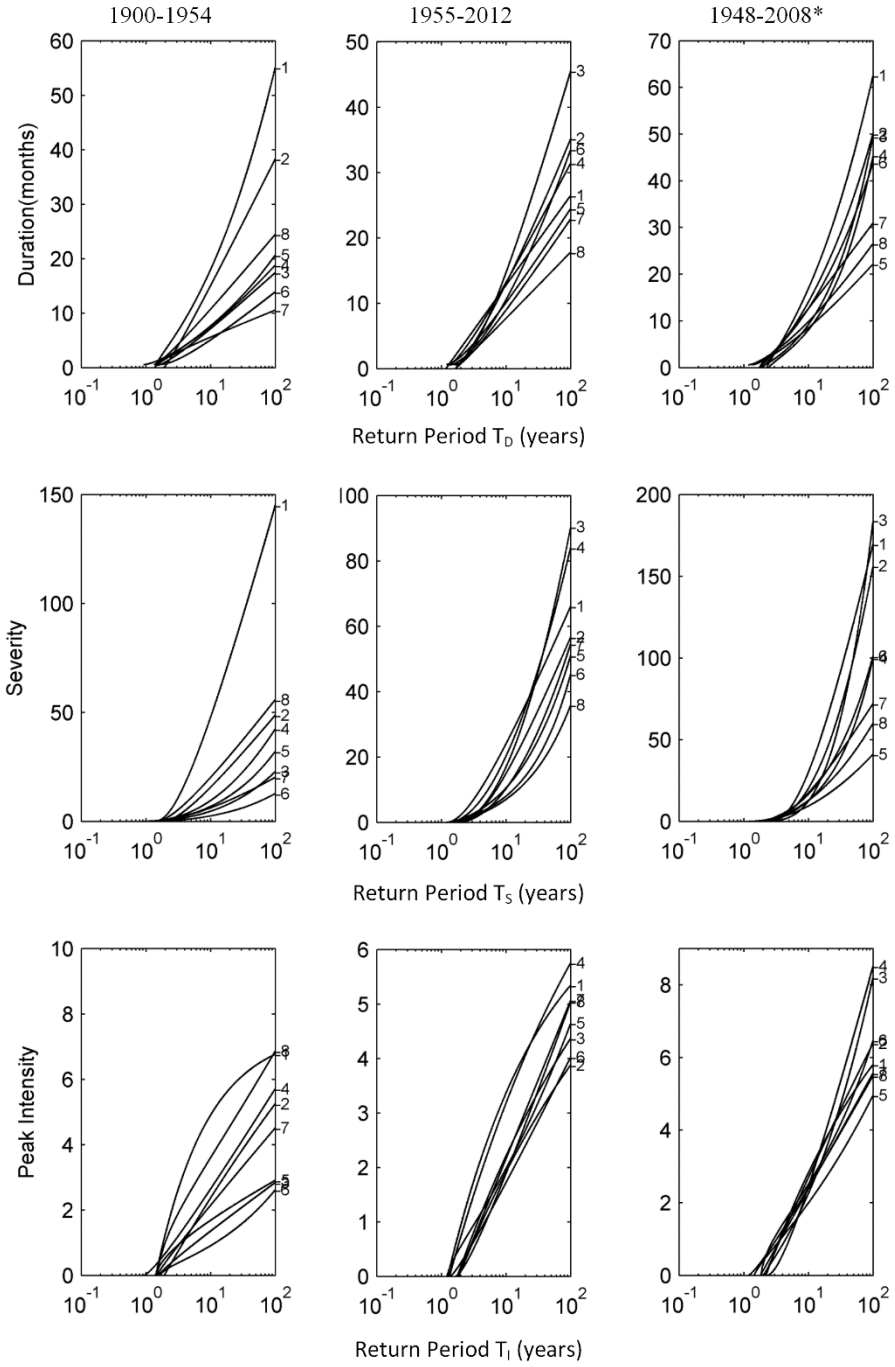


Figure 4.7 Spatial comparison of return period for individual drought variables: D, S and I for dataset of Dai (2013) during two historical periods (1900 to 1954, 1955 to 2012), respectively, and for the dataset of Sheffield et al. (2012) during the period of 1948 to 2008 in eight areas in northern and eastern India.

(2) Bivariate return periods

Using Equation (17) for the pair of D-S, joint bivariate return periods of "D and S", "D and I" are calculated for all areas in the three historical periods, shown as the contour plots in Figure 4.8 and Figure 4.9, respectively. As expected from the definition (see equations 14-17), one can observe that $T_{\cap(DS)}$>T_{S}, T_{D} . This is because the probability of two concurrent occurrences (i.e., S exceeds a certain level and so does D), $P(D \geq d, S \geq s)$, should be smaller than either of the two occurrences $P(D \geq d \; or \; S \geq s)$. For example, for area 1 during 1900-1954: $T_{(D\geq 15 \; and \; S\geq 40)} = 10.9$>$T_{(\mathrm{D}\geq 15)} = 9.1$, $T_{(\mathrm{S}\geq 15)} = 8.6$. Similar results are observed for the pair of D and I.

Return periods of a particular pair of drought variables vary by time period. We take area 1 as an example to compare the joint bivariate return periods in sequential time periods. Bivariate return period $T_{\cap DI}$ (D and I) mostly depends on duration when peak intensity approaches the limit (around 5.8) during 1900-1954, while the pattern of 'return period $T_{\cap DS}$ (D and S) is not subject to such a condition. This is because of the definitions of S and I: for one drought event, there is a bound of the peak intensity (PDSI usually takes a range of -6 to 6); meanwhile severity can be accumulated to be any value theoretically with long-lasting drought events. Drought events based on both short (e.g., the 5-year) and long (the 65-year) bivariate return periods correspond to lower D and I (lower D and S) during 1955-2012 than that during 1900-1954 for $T_{\cap DI}$ ($T_{\cap DS}$), which means drought quantified as bivariate return period of "D and I" ("D and S") became less severe after 1955.

Bivariate return periods with of a particular pair of drought variables also vary by areas. Comparing area 1 and area 3 (Figs. 8 and 9), drought in all aspects had become more severe in area 3 after 1955 than area 1, which can also be concluded from the comparison of univariate return period. During 1900-1954, drought occurred with much shorter D, less I and lower S in

area 3 than area 1; however during 1955-2012, similar drought features were observed in the two areas except for a slightly longer D in area 3.

The contour plots of bivariate joint return periods with “D and S” and those with “D and I”, as shown in Figs. 8 and 9 represent the Duration-Severity-Frequency (D-S-F) curve and Duration-Intensity-Frequency (D-I-F) curve, respectively. These curves provide the probability of occurrence of drought events with a specific duration and severity/intensity, which can be used to 1) identify the so-called design droughts used for planning and design of drought preparation measures (Halwatura et al., 2015), and 2) determine the risk of failure of current local drought preparation measures. When D, S and I change in the same direction from the early to later period and the changes are considerably large (e.g., changes in areas 1 and 3-8), the corresponding D-S-F and D-I-F change in the same direction in all areas except for area 2. In area 2, there was decreasing D and I under all return periods (see Table 4.1 and Figure 4.6), but increasing S under a long return period but decreasing S under a short return period (e.g. 10-year). In addition, the change in magnitudes of D and S under all the return periods in the two periods are relatively small in area 2. The bivariate D-S-F and D-I-F changes did not follow the pattern of univariate characteristics probably due to the correlation between D and S or D and I. For example, the values of D and I as a pair under the 5-year return period during 1955-2012 were higher than those during 1900-1954 (e.g., D= 7.0 and I=0.8 during 1955-2012 vs. D= 6.5 and I =0.8 during 1900-1954). Thus we need to assess the drought characteristics under a bivariate return period beyond the assessment of univariate status.

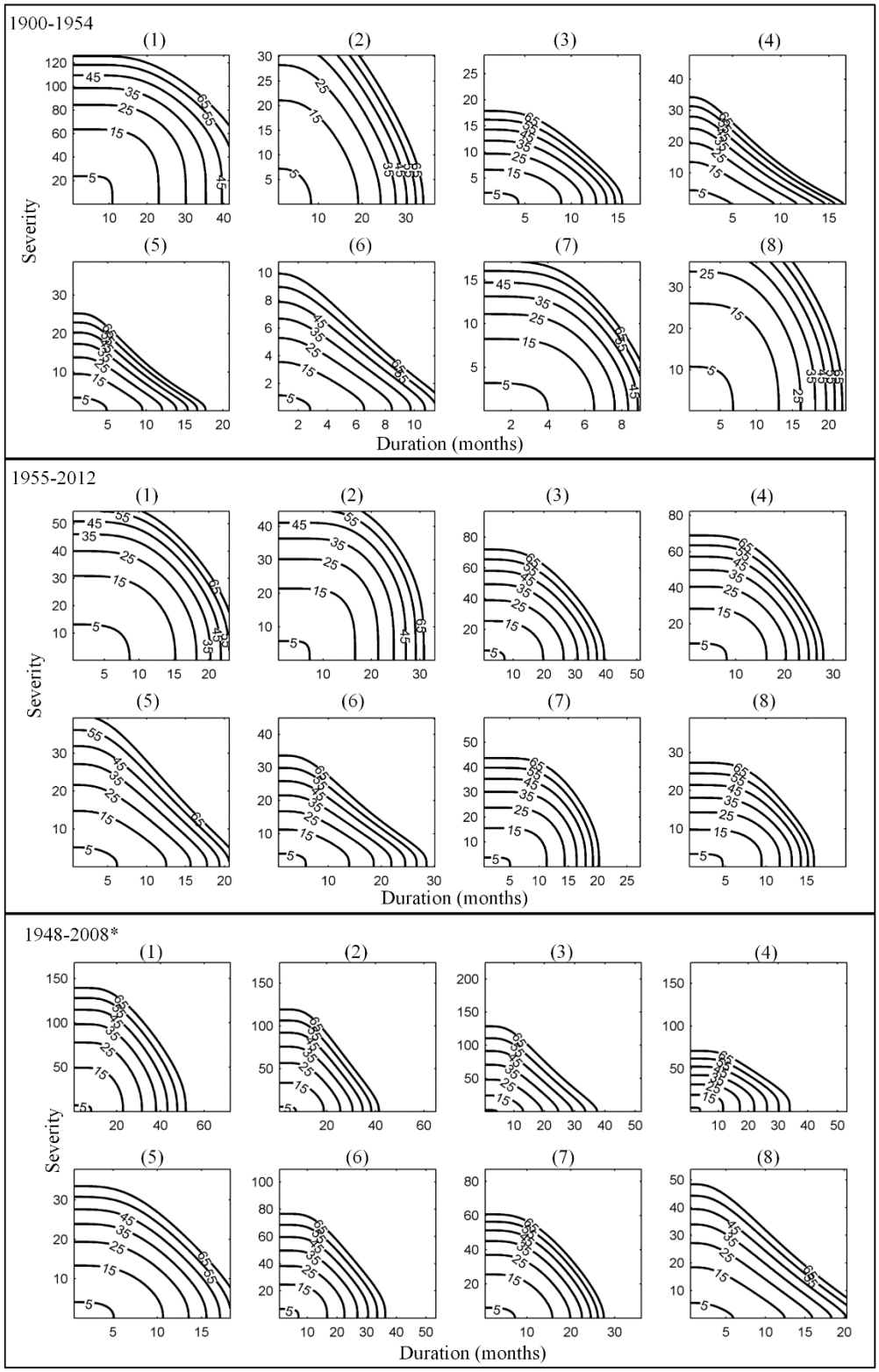


Figure 4.8 Contour plots of bivariate joint return periods "D and S" in eight areas during three periods: 1900-1954 and 1955-2012 of Dai (2013), 1948-2008 of Sheffield et al. (2012).

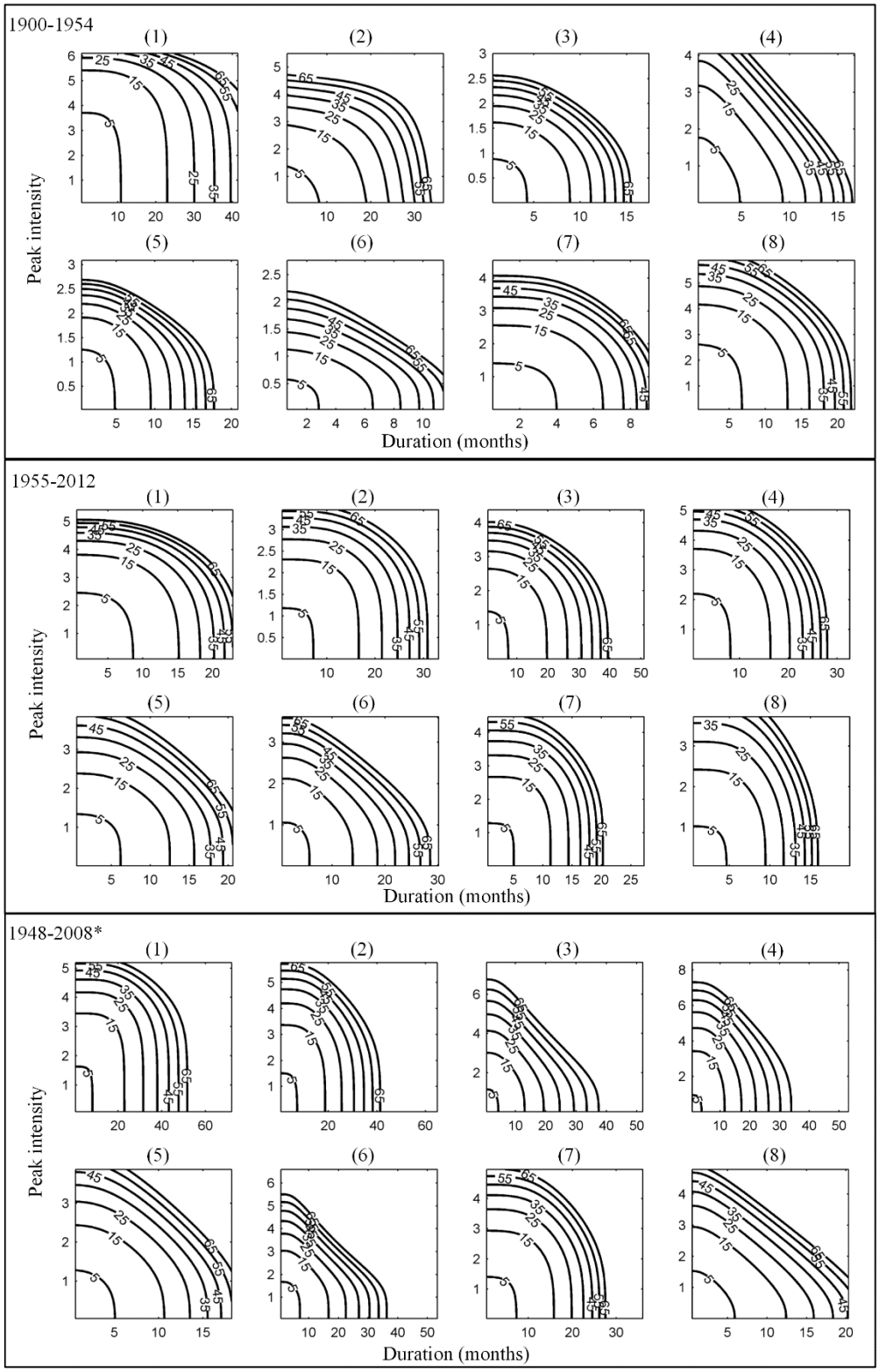


Figure 4.9 Contour plots of bivariate joint return periods "D and I" in eight areas during three periods: 1900-1954 and 1955-2012 of Dai (2013), 1948-2008 of Sheffield et al. (2012).

4.2.4 Discussion and Conclusions

Temporal and spatial comparisons show different changes with severity (S), duration (D) and peak intensity (I) during past periods (1900-1954 and 1955-2012) in northern and eastern India. Considering multiple variables of drought enables assessing drought events and their impacts in a more comprehensive approach than the traditional one that is based on the time series of the mean value of a drought index. It is found that in some areas drought after 1955 can be more harsh with respect to some but not necessarily all characteristics. For example, during period 1955-2012, area 2 experienced a longest D, medium S but lowest I among all eight areas, showing a case of long, least intense drought events cumulating and ending with a moderate level of S. Furthermore, drought frequencies vary over space substantially even in the same region of northern and eastern India. For example, in the same state (Uttar Pradesh), area 2 and area 3 experienced significantly different changes in drought frequencies when comparing the two historical periods; duration and peak intensity are declining in area 2 after 1955, but with an opposite trend in area 3. Generally, droughts in the areas which mainly plant wheat (areas 1 and 2) became shorter and less intensive after 1955 while droughts in the areas with rice as the major crop (areas 3-8) became longer, more severe and more intensive (except for area 8, a coastal area, had shorter, less severe and less intensive droughts after 1955). This spatial change pattern may imply crop pattern change, for example, switching from rice to wheat in areas 3-7. Drought characteristics (D, S and I) can impact agriculture and water resources management in different ways. Generally, highly intensive droughts with even a short or medium duration (seasonal or shorter) can have a large impact on agriculture (e.g., the 2012 drought in the U.S. Midwest, which occurred during the growth period of corn); while, medium intensity droughts with long

durations can have very serious effect on water supply (e.g., the continuous drought in California).

Results also show that D, S and I are correlated and bivariate joint return periods can be identified by copulas. The joint bivariate return periods vary by time and space. In most of the study areas, both the temporal and spatial change patterns of D-S and D-I under bivariate return periods are consistent with those of single characteristics under univariate return periods. For example, in area 1, drought quantified as bivariate return period of "D and I" ("D and S") became less severe after 1955. Also, in area 3, drought was more severe after 1955 than another (e.g., area 1), same as the pattern in terms of three characteristics (D, S and I) (due to the positive correlation among the three drought variables). However, the drought characteristics with a joint bivariate return period can be more complex than those with a univariate return period due to the correlation between individual drought characteristics. In particular, the temporal change pattern of D-S or D-I under various bivariate return periods can be different from that of D, S or I under univariate return periods (e.g., the case of area 2).

However, uncertainties exist with the assessment of multiple drought characteristics as shown by the different results from the two datasets (generated by Dai, 2013 and Sheffield et al., 2012). Note that even larger differences in terms of D, S and I are found with the two datasets than those with PDSI as compared in previous studies (Trenberth et al., 2013). In particular, major differences in D and S under long-term return periods are identified between the two datasets due to different climate forcing inputs (especially regarding uncertainty representation) and model structures (Trenberth et al., 2013).

4.3 Spatial and Temporal Patterns of Drought in the Continental U.S. during the Past Century

In this study, trends in D, S and I levels are analyzed for the Continental U.S. There have been debates on recent trends in droughts around the world. While some studies claim that there has been increasing global droughts since 1950 (Vicente Serrano et al., 2010; Dai, 2013), Sheffield et al. (2012) argued that the increasing trends were due to the methodology adopted by those studies for estimating potential evapotranspiration. For the Continental U.S., however, the results of the studies analyzing drought trends seem to have a better agreement. For example, most studies show increasing dryness in the Western U.S., increasing wetness in the Eastern U.S. (Groisman et al., 2004; Andreadis and Lettenmaier, 2006; Ficklin et al., 2015), and an increase of summer precipitation variability in the Southeastern U.S. (Li et al., 2010; Li et al., 2013). However, according to our knowledge, trends in the characteristics of drought events (D, S and I) with different return periods have not been analyzed for the continental U.S. This paper attempts to fill this knowledge, especially by identifying both the spatial patterns of D, S and I trends and the temporal variation patterns during 1900 to 2012 in the Continental U.S. We use a trend analysis method based on the overlapping of moving windows proposed by Liu et al. (2014) to identify continuous changes in drought characteristics, as detailed in the following section.

4.3.1 Data Sources and Methodology

In this part of the study, we use PDSI for drought identification and quantification, since PDSI is likely the most popular regional drought index for monitoring droughts and for assessing agricultural impacts (Mishra and Singh, 2010). We use two datasets of the monthly time series of PDSI 1) by Dai (2013) (PDSI_d) with 2.5 degree spatial resolution from 1900 to 2012, and with 140 sites located in the Continental US, and 2) by Sheffield et al. (2012) (PDSI_s) with 1 degree spatial resolution from 1948-2008, with 826 sites located in the Continental US. These two

datasets have the following advantages: a) both the datasets use Penman-Monteith equation for the calculation of Potential Evapotranspiration (PET) (which has been found to be more physically realistic and based on more diverse input data than the Thornthwaite equation) (Dai, 2013; Sheffield et al., 2012); b) the datasets use self-calibration method of PDSI calculation (Wells et al., 2004), in which PDSI is calibrated using local conditions ; c) the two datasets have been recently used for studying drought trends at global scale (Sheffield (2013) and Sheffield et al (2012)). In addition, the PDSI_d dataset is the longest available PDSI record and is hence ideal for long term temporal analysis of droughts. It should be noted that several other historical drought index datasets are available with even finer resolutions. However, because the aim of this study is to analyze long term trends in drought characteristics for a large scale region, both adequate length and spatial coverage of the data are important, and the two selected datasets satisfy these requirements. To focus on relatively severe drought events, the critical level of PDSI is set to -1 (corresponding to "mild drought" in PDSI identification) (Madadgar and Moradkhani, 2013) in this study, which means that only the events with a PDSI less than -1 are analyzed for the drought trend assessment. Drought characteristics (D, S, I, and T) are calculated using monthly PDSI time series.

In this study, seven commonly used uni-variate probability distributions are used as the candidates to fit the distributions for the D, S and I of drought events: Exponential, Gamma, Weibull, Log-normal, Logistic, Generalized Pareto and Generalized Extreme Value. The Kolmogorov-Smirnov (KS) test (Marsaglia et al., 2003) is used to select the best fitted distribution for each drought characteristic.

Some recent studies analyzed drought characteristics separately with long and short return periods (e.g., Ganguli and Ganguly, 2015; Ge et al., 2016; Liu et al., 2016) and compared

drought characteristics under a given return period in different time periods (e.g., 1900-1954 vs. 1955-2012 in Ge et al. 2016; 1926-1969 vs. 1970-2013 and 1950-1979 vs. 1980-2009 in Ganguli and Ganguly, 2015). The use of two segmented periods provides a comparison of only two values which are representative of drought conditions in that period. However, the temporal changes in drought characteristics can be very complicated: while there may be long term trends in drought characteristics, there can also be significant variability at shorter time scales. Thus, using two segmented periods may not provide information about the continuous change process of drought characteristics. Further, the results with discrete time periods are sensitive to the choice of periods selected for comparison. The results are found to vary with the periods selected for comparison (Ganguli and Ganguly 2015). In order to assess the temporal variation of drought characteristics, a moving window approach is applied by generating sequential 40-year moving windows (e.g., 1900-1939, 1901-1940, …, 1973-2012, 74 windows for the PDSI_d dataset (2013); 1948-1987… 1969-2008, 22 windows for the PDSI_s dataset), following the approach proposed by Liu et al. (2014). The levels of D, S, and I under a certain return period (e.g., 5-year, 10-year, 25-year, 50-year, and 100-year) are calculated considering drought events with expected drought inter-arrival times (Shiau and Shen, 2001; Shiau 2006) and the corresponding selected best fit distributions for D, S and I in each window. Hence, 74 values corresponding to the 74 windows using the PDSI_d dataset and 22 values corresponding to the 22 windows using the PDSI_s dataset for each of the drought characteristics D, S and I under a certain return period are obtained at each site.

Short time Fourier transform (STFT) (Gabor, 1946) is applied to the time series of magnitude of drought characteristics under different return periods obtained from the PDSI_d dataset to characterize and classify temporal changes in drought characteristics. STFT involves

the application of Fourier transform to overlapping moving windows. STFT is considered in this study as the time series of drought characteristics may not be stationary during the entire period of study. Note that we only apply the classification method to the PDSI_d dataset to identify the long-term change patterns, as; the PDSI_s dataset is not long enough for the purpose.

Lastly, 25 windows (i.e., 1949-1988… 1973-2012) based on the PDSI_d dataset and 22 windows of the PDSI_s dataset (i.e., 1948-1987… 1969-2008) are chosen to conduct trend analysis for D, S and I under different return periods at each site. This is because rapid increases in global surface temperatures have been observed due to global warming (Dai, 2013) during 1949-2012, and the analysis of drought trends during this period may reveal the impact of climate change on droughts. The Trend Free Pre-whitening procedure (TFPW) (i.e., the R package zyp provided by Bronaugh and Werner, 2013) is used to detect whether there is a positive or negative trend (significance level is chosen as $\alpha = 0.05$ to determine if a trend is statistically significant) associated with the time series of D, S or I based on moving-windows.

4.3.2 Results and Discussions

4.3.2.1 Spatial distribution of D, S and I

The spatial distribution of D, S and I levels under a short (10-year) and a long (50-year) return period are shown in Figure 4.10. The spatial distributions have been plotted for three selected 40-year windows (1900-1939, 1940-1979, and 1973-2012, the 1^{st} , 41^{st}, and 74^{th} window, respectively) to investigate the temporal variation of the spatial distributions. The values of D, S and I under 10-year and 50-year return periods have been calculated based on the fitted distributions for the corresponding drought characteristics in the three windows (a site in Northern California is shown as an example in Figure 4.11 in Appendix C to illustrate the goodness of fit).

As seen in Figure 4.10, the spatial distributions of D, S and I for the 10-year return period and the 50-year return period of the same time period are quite different. For example, during 1973-2012, the severity under the 10-year return period in northwestern coastal areas is amongst the highest in the continental U.S.; however, under the 50-year return period, during the same time period, the severity values are not so high in this region as compared to other regions in the U.S. Similar examples of different spatial distributions of drought characteristics under different return periods during a same time period include intensity in the Midwest during 1940-1979, severity in the Western U.S. during 1940-1979 and duration in California during 1973-2012.

As D, S and I are highly correlated, the spatial patterns under the same return period and during the same time period are similar. However, in some areas, the spatial patterns of I are quite different from those of S and D during the same time period. For example, during 1900-1939, S and D values in the northwestern coastal areas are amongst the highest in the U.S., but the intensity values are low in this area during the same period under both return periods. Similar observations can be made for the Great Lakes region during 1940-1979.

Additionally, it can be observed from Figure 4.10 that the spatial patterns of drought characteristics vary with time. For most sites across the U.S., the values of D, S and I under both the 10-year return period and the 50-year return period are lower in the 1940-1979 period when compared to the other two periods, especially for I. The spatial patterns of D, S and I are quite similar in the periods 1900-1939 and 1979-2012 (high D and S in Northwest, Midwest and Southeast areas); the difference consists of higher D and S values in the Great Lakes and the central U.S. regions during 1900-1939.

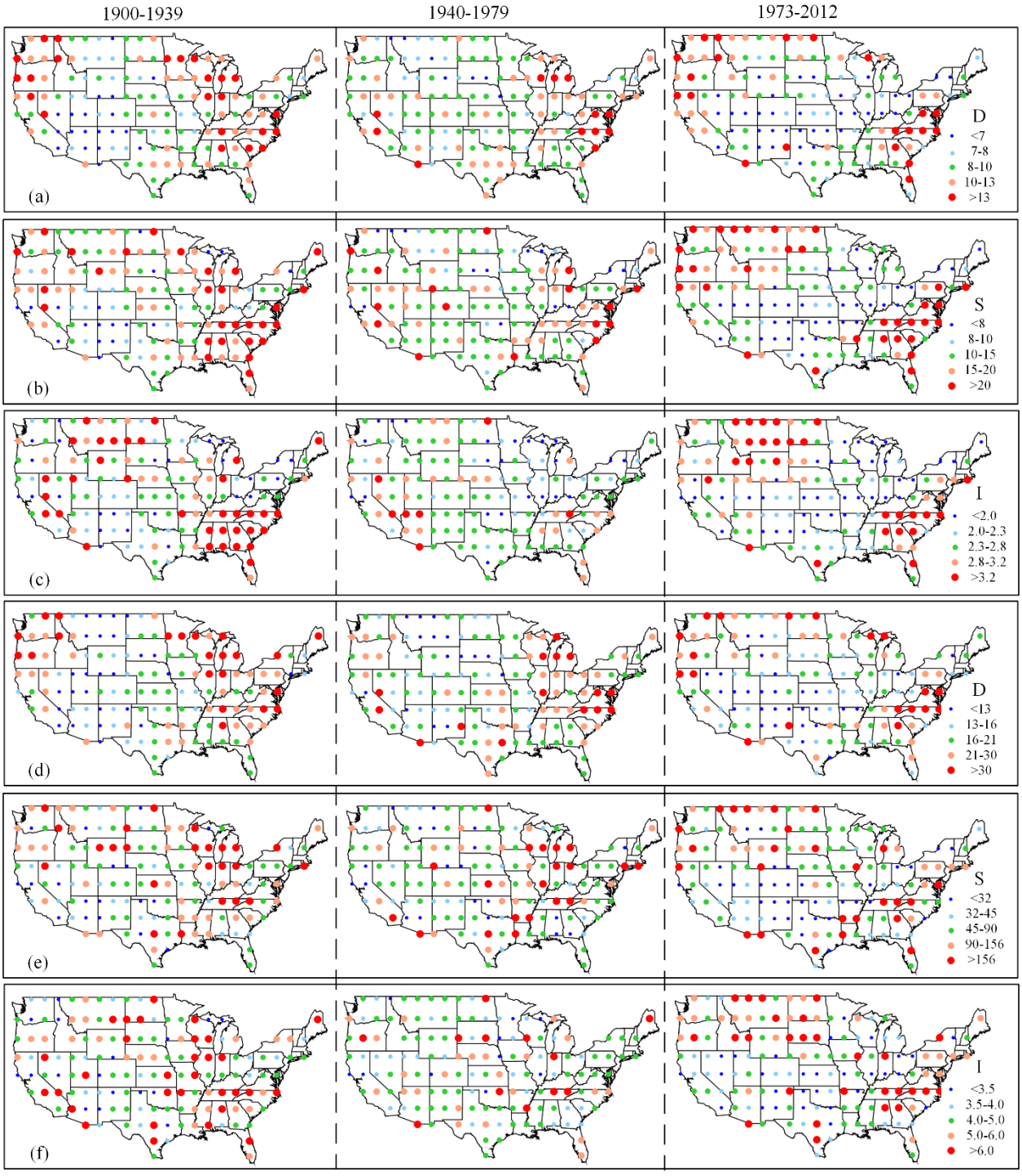


Figure 4.10 Spatial distribution of magnitudes of (a) D, (b) S and (c) I under the 10-year return period and (d) D, (e) S and (f) I under the 50-year return period in three windows: 1900-1939 (1st), 1940-1979 (41st) and 1973-2012 (74th).

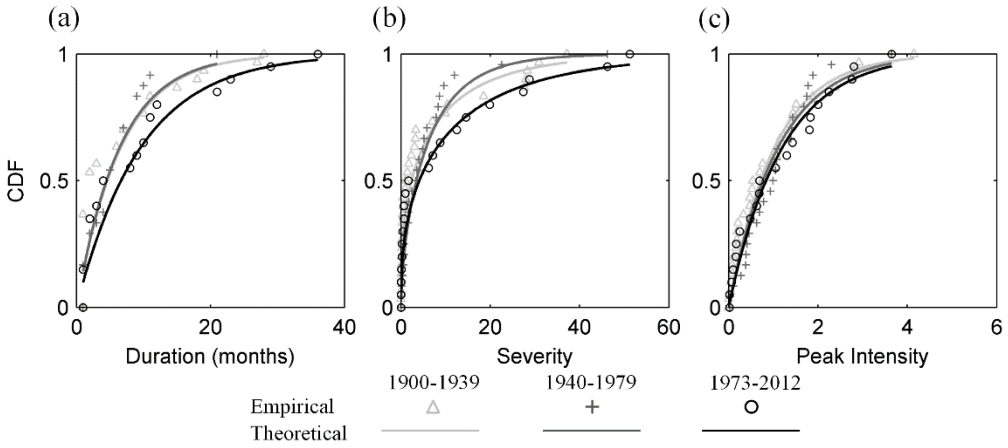


Figure 4.11 Comparison of empirical and theoretical CDF of marginal distributions of D, S, and I in three sub-windows: 1900-1939 (1st), 1940-1979 (41st) and 1973-2012 (74th) in the example site of Northern California.

4.3.2.2 Classification of change patterns

The spatial distributions of D, S and I in the selected three periods shown in Figure 4.10 provide a comparison of spatial patterns of the drought characteristics in three discrete time periods. In order to investigate the continuous temporal variation of drought characteristics we generate a set of time series of D, S and I for each site under different return periods calculated using the 74 overlapping moving windows. STFT was applied to the time series of the magnitudes of the drought characteristics for different return periods to obtain the power density spectrum of the time series at each site. All locations show the same dominant low frequency component although the power of the dominant frequency varies from site to site. The sites were then classified into two categories based on the power of the dominant frequency in the power spectrum: pattern I (persistent) with the dominant frequency power more than the median value and pattern II (transient) with the dominant frequency power less than the median calculated value across all sites. Classification of sites based on D, S and I time-series under different return periods provided quite similar results. Figures 4.11(a) and 4.11(b) show two typical examples of Pattern I temporal variation for a site located in northern California and North Dakota

respectively. As seen in the figure, Pattern I temporal variation can be characterized by significant changes in the magnitude of the drought characteristics over long periods of time for a certain return period (i.e., frequency of a drought event). The figure also shows that in northern California, all drought characteristics D, S and I show a significant increasing trend under almost all return periods approximately from 1980, especially under longer return periods; while in North Dakota, all drought characteristics D, S and I show a significant decreasing trend after 1990s. Figure 4.11(c) shows a typical example of Pattern II, for a site located in northern Mississippi. As can be seen, Pattern II temporal variation is characterized by short term changes in the magnitude of drought characteristics, and a clear trend in D, S and I levels is not evident under any return period in northern Mississippi.

Furthermore, by analyzing the spatial distribution of sites showing Pattern I and Pattern II temporal variations, we find that sites with Pattern I temporal variation are mostly located in the West coast, the Great Plains, the Northeast and a few sites in the Southeast, whereas sites showing Pattern II temporal variation are mostly located in the interior sites of the continental U.S (Figure 4.13). Regions with persistent drought conditions (Pattern I) are likely to suffer more damage as the drought conditions prevail for a longer period of time.

Droughts in Western U.S. and parts of the northern Great Plains, where Pattern I temporal variation is mostly identified, are influenced by Pacific sea surface temperature (SST) anomalies (Cook et al., 2004; Schubert et al., 2004; Weiss et al., 2009). Studies have shown that warming of tropical Pacific SSTs produce extreme drought events in the region. For example, droughts of the 1930s (Dust Bowl), 1950s, 1980 and 1988 led to large changes in magnitudes of drought characteristics resulting in Pattern I temporal change patterns. Another region which shows Pattern I temporal change is Eastern U.S., including Virginia, South Carolina and North

Carolina. Studies have shown that droughts in this region are influenced by tropical Atlantic SSTs (Enfield, 1996; Enfield et al., 2001; McCabe et al., 2004). Anomalous warm Atlantic SSTs produce extreme droughts in the region, which are responsible for Pattern I temporal variation.

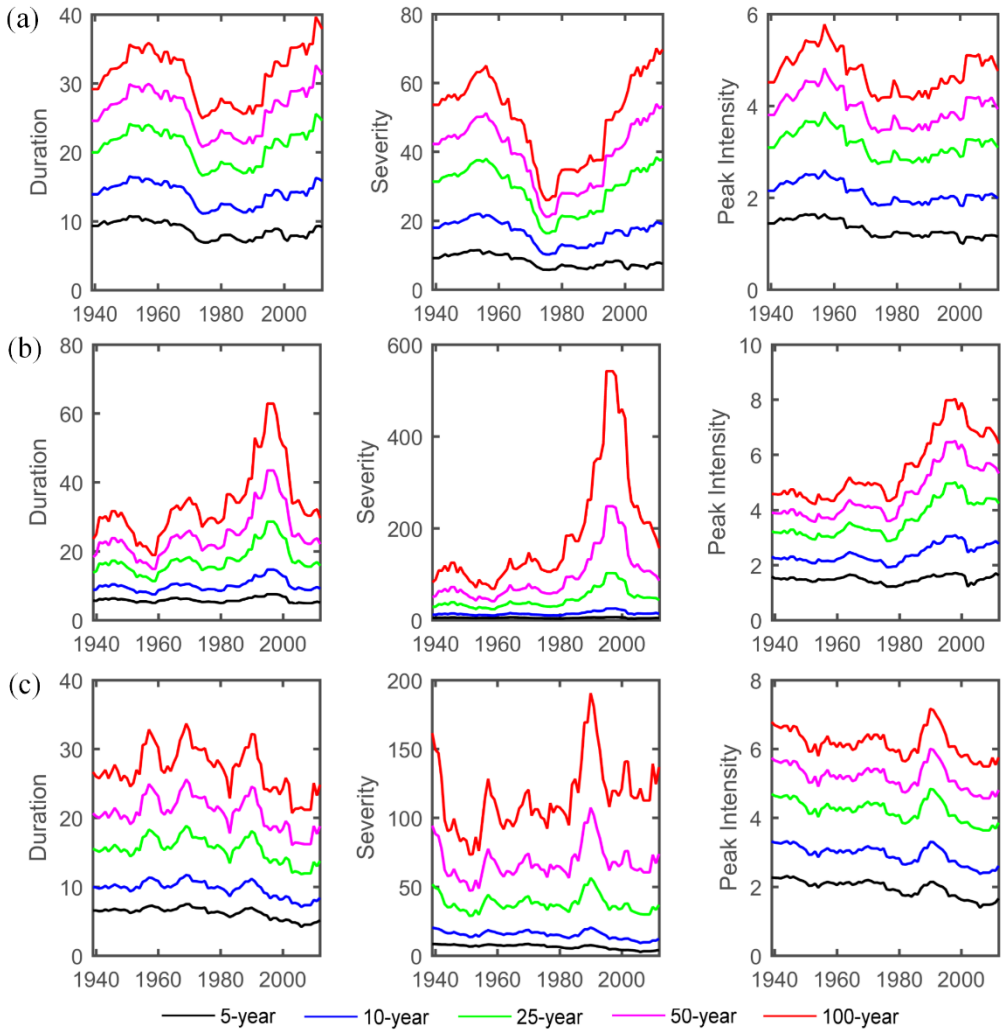


Figure 4.12 Levels of D, S, I under different uni-variable return periods in sequential moving windows (a) and (b) in two example sites of pattern I in Northern California (41.25, -121.25) and North Dakota (46.25, -98.75) and (c) in an example site of pattern II in Northern Mississippi (33.75, -88.75).

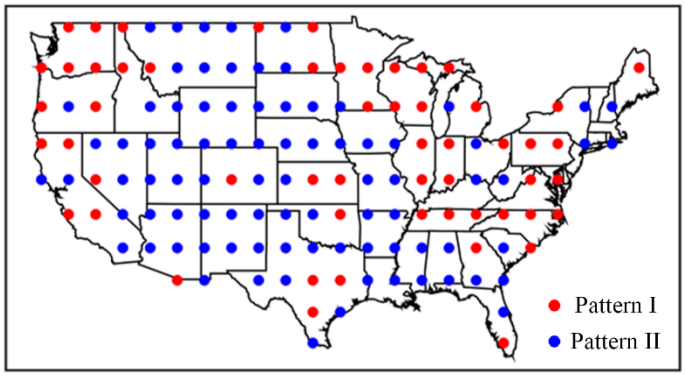


Figure 4.13 Classification of change patterns of return periods in the continental US. Pattern I (red): persistent drought; pattern II (blue): higher variability in drought index.

4.3.2.3 Trend detection of drought characteristics under different return periods

The examples in Figure 4.12 show that the temporal changes in drought characteristics are not consistent over the whole period. In order to assess the recent trends of D, S and I levels, we chose the last 25 windows (1949-1988, …, 1973-2012) of the PDSI_d dataset since the global surface temperature has been increasing from around the 1950s, and all the 22 windows of the PDSI_s dataset (1948-1987… 1969-2008); thus the results from the two datasets are comparable in terms of the time frame. Figure 4.14 shows the locations with significant or insignificant (with p=0.05) positive or negative trends for drought characteristics D, S and I, which is calculated using the TFPW procedure (Bronaugh and Werner, 2013). The results of PDSI-s using cpc1.0 precipitation dataset are shown in Figure 4.14; those using using cru3.10, gpcc4 and willmott2.01 precipitation dataset are shown in Figure 4.15-4.17. The trend analysis of D, S and I is conducted separately for three different return periods (10-, 25-, 100-year). Figure 4.14 displays a spatial pattern crossing the sites with increasing or decreasing trends of a drought characteristic. Generally, most sites with an increasing trend in D, S and I are located in the northwest, the east, and the southwest under the three different considered return periods based on the PDSI_d dataset (2013) (Figure 4.14(a)). The trends of drought characteristics

calculated using the two datasets appear to be quite similar, except in the Midwest where more increasing trends in D, S and I are found when using the PDSI_s dataset, especially for S and I. With different precipitation data input sources, the trends at particular sites could be different, revealing the impact of uncertainty associated with PDSI calculation (Figure 4.14(b) and Figure 4.15-4.17). However, overall the large scale spatial patterns of trends are similar. Andreadis and Lettenmaier (2006) also found an increasing trend in drought duration and severity in western coastal areas and in some southeastern areas, even though their analysis used soil moisture as an indicator of drought. In addition, it can be seen that most sites with increasing drought duration or severity (around 71%) at all return periods show Pattern I temporal variation in drought characteristics. The spatial distribution of locations with increasing or decreasing trends is similar for D, S and I due to the high correlation between these drought characteristics (Figure 4.14). However, some differences do exist, especially for I. For example, in the PDSI_d dataset (Figure 4.14(a)) there are more sites with increasing I in the northwest in comparison to D and S (e.g., Montana). This confirms that trend analysis applied for multiple drought characteristics (D, S and I) reveals more information than when applying trend analysis directly to a drought index time series. The trend analysis applied to D, S, and I separately can conclude different results when compared to trend analysis applied directly to drought index. For example, the areas with increasing or decreasing drought characteristics are different from the areas with increasing or decreasing PDSI found by Ficklin et al. (2015). The results from the PDSI_s dataset (Figure 4.14(b)) supports the finding that there can be different trends in drought characteristics at one site, e.g., the results of the PDSI_S dataset shows that in Montana there are a larger number of sites with increasing I as compared to sites with increasing D and S. Further, PDSI_s dataset with fine resolution reveals that there are quite a few regions showing different trends in different

drought characteristics, e.g. Iowa shows increasing trends in S and I but a decreasing trend D and Pennsylvania shows no significant trend for D but increasing trends for S and I.

For most of the locations, trends in drought characteristics are similar under different return periods, but for some sites the trends under short return periods are different from those under longer return periods examples include D in Idaho, Nebraska and Arizona; S in the southern California, Alabama and South Dakota; I in California and Utah when the PDSI_d dataset (Figure 4.14(a)) is considered. The use of the fine resolution PDSI_s dataset further corroborates these findings by revealing spatial extent of the regions identified using the PDSI_d dataset which have different trends under different return periods e.g. in the PDSI_d dataset we find certain sites in southern California, where there are opposite trends in S under different return periods; for the same region the PDSI_s dataset provides more detailed information by showing the regions represented by the sites of the PDSI_d dataset which have different trends in S under different return periods. It should be noted that the results based on continuous moving windows are quite different from those of previous studies which have used discrete time periods for the comparison of drought characteristics. For example, Ganguli and Ganguly (2015) showed different results when the periods for comparison were chosen differently. They found that there were more sites with larger severity during 1970-2013 compared to 1926-1969 in central, northeastern, northwestern, southern, and western areas; however, there were more sites with smaller severity during 1980-2009 when compared to 1950-1979 in all the same regions except for the central region. The method of continuous moving windows used in this study not only provides a clearer picture of the temporal changes in drought characteristics but also avoids the problem of choosing the discrete time periods for the comparison of drought characteristics. This avoids the possibility of different results due to the choice of the periods for comparison.

The results of the trend analysis in our study can be related to previous studies, which have tried to explore the causes of drought patterns observed in the U.S. In this study, we found that most of the sites showing an increasing trend in D, S and I also show Pattern I temporal variation of drought characteristics in the last century. The regions showing an increasing trend in drought characteristics include the northwest, southwest and northern Great Plains. As mentioned earlier, droughts in this region are significantly influenced by Pacific SSTs. Previous studies (Cook et al., 2004; MacDonald, 2008) have found that during 800-1300 AD, warmer tropical Pacific SSTs produced extreme droughts with a duration 5-8 years using tree ring analysis. Therefore, gradual warming of the tropical Pacific due to global warming might be a cause of the increasing trend in drought characteristics observed in this region. In addition, some sites in the southeast (including Virginia, South Carolina and North Carolina), where droughts are influenced by Atlantic SSTs, also show an increasing trend in drought characteristics. Studies have shown that in this region, gradual warming of tropical Atlantic SSTs is leading to higher variability in summer precipitation in the last 30 years, causing more droughts in the region (Enfield, 1996; McCabe et al., 2004; Wang et al., 2010). An increasing trend in drought characteristics has also been observed in North Eastern states, but unlike other regions with an increasing trend, studies have found that droughts in this region are mainly caused by regional climatic processes (Hayhoe et al., 2007). Thus, the sites where droughts are strongly influenced by tropical Atlantic and Pacific SSTs are showing an increasing trend in drought characteristics possibly due to global warming.

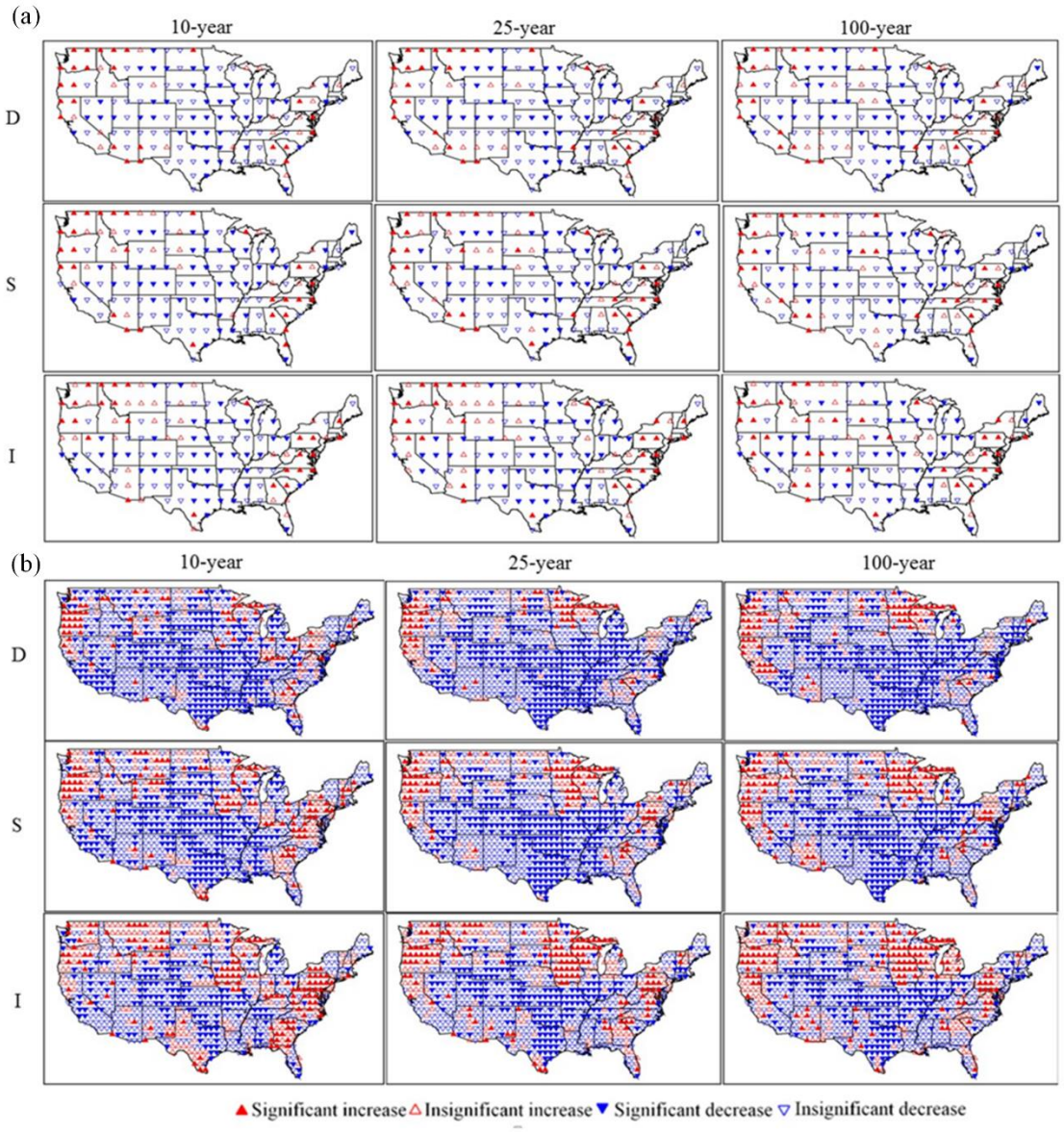


Figure 4.14 Trends of levels of D, S and I under different uni-variable return periods (10-, 25-, 100-year) in (a) selected last 25 windows of the PDSI_d dataset (2013) and the total 22 windows of the PDSI_s dataset (2012) using cpc1.0 precipitation dataset.

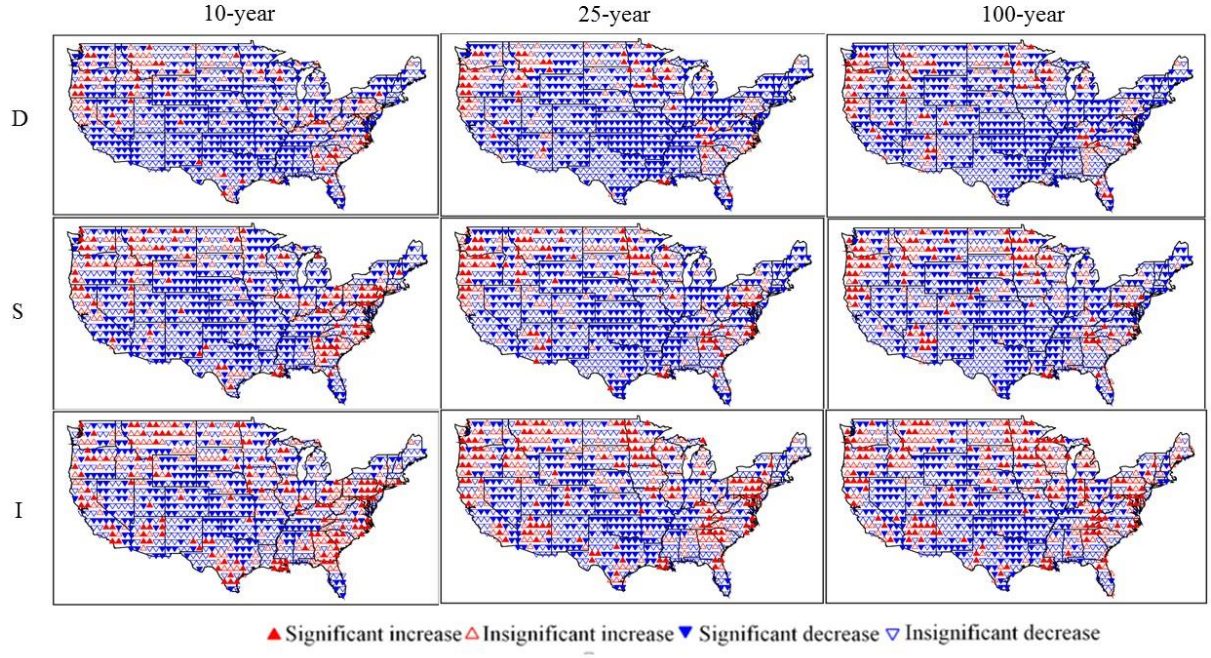


Figure 4.15 Trends of levels of D, S and I under different uni-variable return periods (10-, 25-, 100-year) in the study period of 1948-2008 using Princeton dataset based on CRU precipitation input.

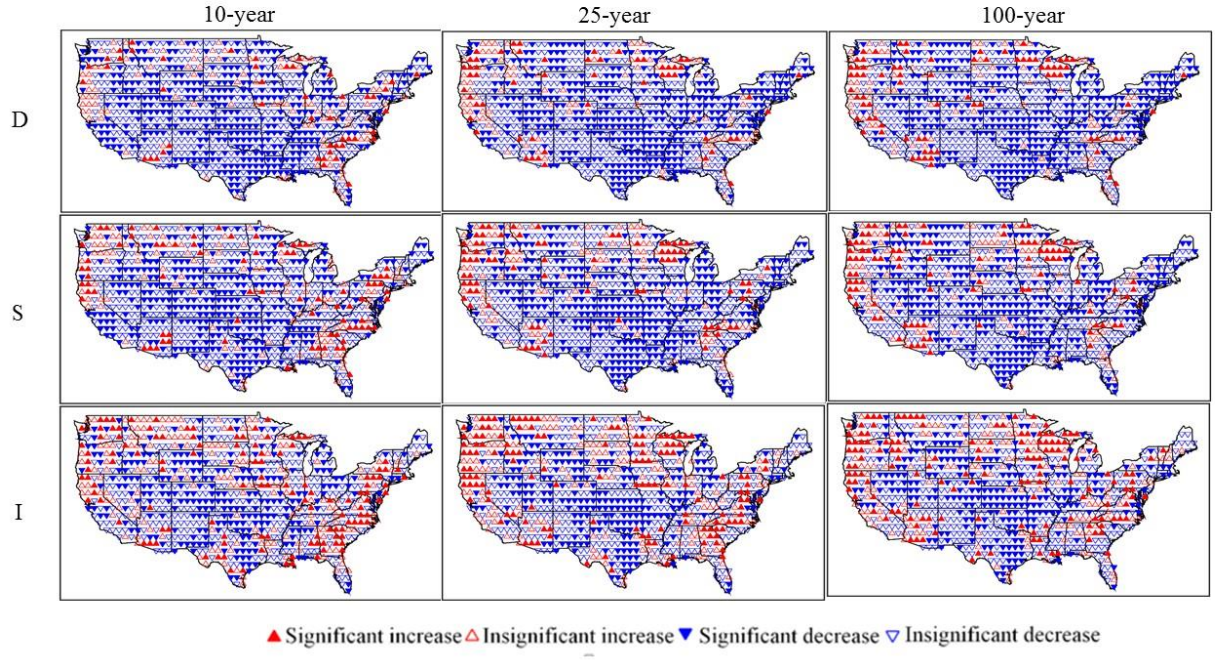


Figure 4.16 Trends of levels of D, S and I under different uni-variable return periods (10-, 25-, 100-year) in the study period of 1948-2008 using Princeton dataset based on GPCC precipitation input.

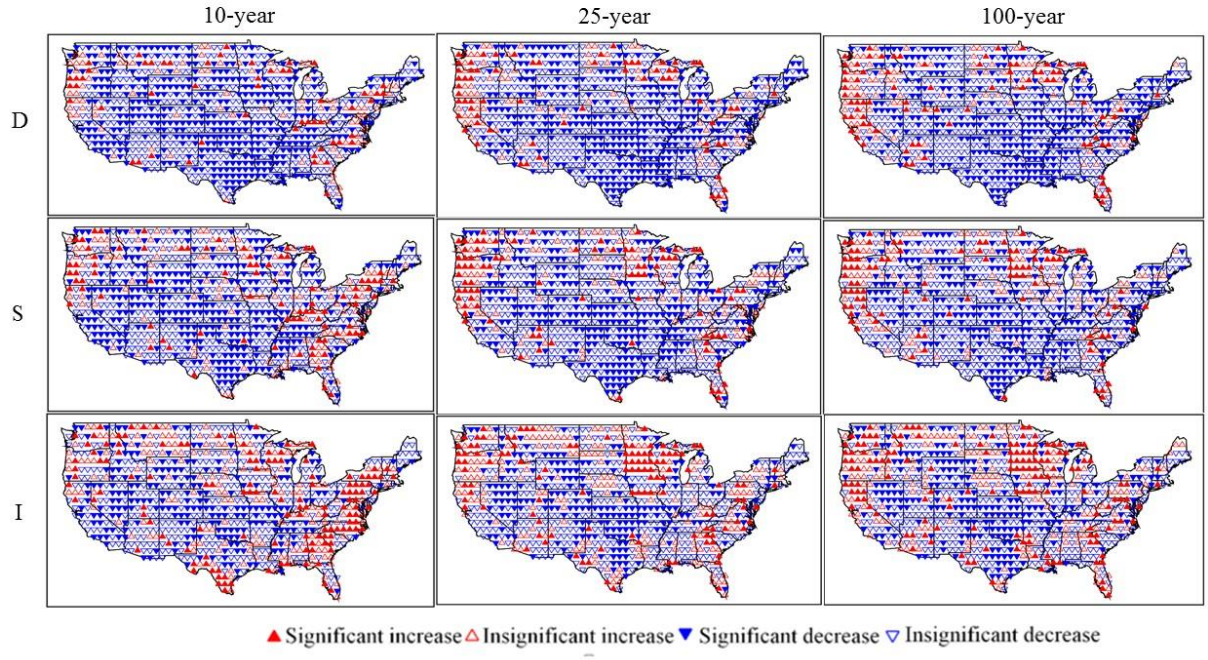


Figure 4.17 Trends of levels of D, S and I under different uni-variable return periods (10-, 25-, 100-year) in the study period of 1948-2008 using Princeton dataset based on Wilmott precipitation input.

4.3.3 Conclusions

This study investigates the spatial and temporal patterns of drought characteristics D, S and I under different return periods during 1900 to 2012 in the Continental U.S using two PDSI datasets: PDSI_d with a longer duration but coarse spatial resolution and PDSI_s with finer resolution and shorter duration. The findings show that Western U.S., the Great Plains, and Eastern U.S. experienced large variations (Pattern I) in the magnitudes of drought characteristics over long periods of time under all the different return periods considered, especially long-term ones (i.e., large variation in extremely severe droughts). The large variations are caused by the occurrence of extreme drought events due to the influence of tropical Pacific and Atlantic SSTs. The interior sites of the continental U.S. show short term variations in magnitude of drought characteristics (Pattern II), indicating transient droughts in this region. It is also found that spatial

distributions of D, S and I for short-term return periods (e.g., 10-year) and long-term return periods (e.g., 50-year) in the same time period can be quite different. Trend analysis shows that D, S and I of droughts under the various return periods are increasing in most of the Pattern I locations (northwest, southwest, northern Great Plains and parts of southeast). The large scale spatial patterns of trends are found to be similar for the two datasets used for almost all regions except the Midwest. Trends in drought characteristics at long and short return periods are similar for most sites; however, at some places, these trends are not the same, showing that trends in extreme and mild drought events can be different. This finding has been found to be consistent in results on drought trends since the 1950's. The classification of change patterns of drought reveals the areas with persistent drought conditions (Pattern I), which may cause more damage to human and natural systems.

4.4 Summary

The spatial and temporal patterns of drought characteristics under different return periods show differences in regions.

The regional analysis is expected to provide important information for selecting appropriate drought mitigation strategies for a particular sector (e.g., agriculture, water supply, ecosystem habitats) of a particular region. The natural (e.g., hydrology) and socioeconomic (e.g., agricultural) impact of a drought event depends on a combined level of D, S and I or one of these characteristics as a dominating factor under some cases. Thus understanding the combined impacts, as well as the dominating impact of one aspect of drought characteristics, is important for drought risk management. The drought frequency analysis based on multiple characteristics of drought events, as presented in this paper, will then provide implications to mitigate the effect of a drought event to a particular natural process or a socioeconomic activity. Especially, the

assessment of expected changes in multiple, correlated drought characteristics is informative for adaptations toward more effective drought management.

4.5 References

AQUASTAT (2016), FAO's global information system on water and agriculture: Food and Agriculture Organization of the United Nations, Aquastat. http://www.fao.org/nr/aquastat.

Andreadis, K. M., Lettenmaier, D. P. (2006), Trends in 20th century drought over the continental United States, *Geophys. Res. Lett. 33*, L10403, doi:10.1029/2006GL025711.

Asquith, W.H. (2015) lmomco---L-moments, censored L-moments, trimmed L-moments, L-comoments, and many distributions. R package version 2.1.4, http://www.cran.r-project.org/package=lmomco, Texas Tech University, Lubbock, Texas.

Bronaugh, D., Werner, A. (2013), Package 'zyp'. CRAN Repository.

Burke, E. J., Brown, S. J., Christidis, N. (2006), Modelling the Recent Evolution of Global Drought and Projections for the Twenty-First Century with the Hadley Centre Climate Model, *J. Hydrometeorol. 7*, 1113–1125.

Carbone, J.G., Dow, K. (2005), Water resource management and drought forecasts in South Carolina, *J. Am. Water Resour. Assoc. 41*(1), 145-155.

Chen, L., Singh, V.P., Guo, S., Mishra, A.K., Guo J. (2013), Drought analysis using copulas, *J. Hydrol. Eng. 18*(7), 797-808.

Clausen, B., Pearson, C. P. (1995), Regional frequency analysis of annual maximum stream flow drought, *J. Hydrol. 173*, 111-130.

Cook, B. I., Ault, T. R., Smerdon, J. E. (2015), Unprecedented 21st-century drought risk in the American Southwest and Central Plains, *Sci. Adv. 1*(1), e1400082, doi:10.1126/sciadv.1400082.

Cook, E. R, Woodhouse, C. A., Eakin, C. M., Meko, D. M., Stahle, D. W. (2004), Long-term aridity changes in the Western United States, *Science 306,* 1015–1018.

Dai, A. (2011), Characteristics and trends in various forms of the Palmer Drought Severity Index (PDSI) during 1900-2008, *J. Geophys. Res. 116*, D12115.

Dai, A. (2013), Increasing drought under global warming in observations and models, *Nat. Clim. Change 3,* 52-58.

De Michele, C., Salvadori, G., Canossi, M., Petaccia, A., Rosso, R. (2005), Bivariate statistical approach to check adequacy of dam spillway, *J. Hydrol. Eng. 10*(1), 50-57.

De Michele, C., Salvadori, G., Vezzoli, R., Pecora, S. (2013), Multivariate assessment of droughts: frequency analysis and dynamic return period, *Water Resour. Res. 49* (10), 6985-6994.

Dracup, J. A., Lee, K. S., Paulson, E. G. (1980), On the definition of droughts, *Water Resour. Res. 16*(2), 297–302.

Enfield, D. B. (1996), Relationships of inter-American rainfall to tropical Atlantic and Pacific SST variability, *Geophys. Res. Lett. 23*, 3305–3308.

Enfield, D. B., Mestas-Nunez, A. M., Trimble, P. J. (2001), The Atlantic Multidecadal Oscillation and its relation to rainfall and river flows in the continental U.S. *Geophys. Res. Lett. 28*, 2077–2088.

Ficklin, D. L., Maxwell, J. T., Letsinger, S. L., Gholizadeh, H. (2015), A climatic deconstruction of recent drought trends in the United States, *Environ. Res. Lett. 10*(4), 044009.

Gabor, D. (1946), Theory of Communication. *J. IEEE 93*, 429-457.

Ganguli, P., Reddy, M.J. (2012), Risk Assessment of Droughts in Gujarat Using Bivariate Copulas, *Water Resour. Manag. 26*(11), 3301-3327.

Ganguli, P., Reddy, M.J. (2013), Spatio-temporal analysis and derivation of copula-based intensity-area-frequency curves for droughts in western Rajasthan (India), *Stoch. Environ. Res. Risk Assess. 27*, 1975-1989.

Ganguli, P., Ganguly, A. R. (2015), Space-time Trends in U.S. Meteorological Droughts, arXiv:1512.08526.

Genest, C., Rémillard, B., Beaudoin, D. (2009), Goodness-of-fit tests for copulas: A review and a power study Insurance, *Math. Econ. 44*, 199-213.

Groisman, P. Y., Knight, R. W. , Karl, T. R., Easterling, D. E., Sun, B., Lawrimore, J. H. (2004), Contemporary changes of the hydrological cycle over the contiguous United States: Trends derived from in situ observations, *J. Hydrometeorol. 5*, 64–85.

Hao, Z., AghaKouchak, A., Nakhjiri, N., Farahmand, A. (2014), Global integrated drought monitoring and prediction system, *Sci. Data 1*, 140001.

Halwatura, D., Lechner, A.M., Arnold, S. (2015), Drought severity–duration–frequency curves: a foundation for risk assessment and planning tool for ecosystem establishment in post-mining landscapes, *Hydrol. Earth Syst. Sci. 19*, 1069-1091.

Hayhoe, K., Wake, C. P., Huntington, T. G., Luo, L., Schwartz, M. D., Sheffield, J., Wolfe, D. (2007), Past and future changes in climate and hydrological indicators in the US Northeast, *Clim. Dynam. 28*(4), 381-407, doi:10.1007/s00382-006-0187-8.

Hisdal, H., Stahl, K., Tallaksen, L., Demuth, S. (2001), Have streamflow droughts in Europe become more severe or frequent? *Int. J. Climatol. 21*, 317-333.

Hofert, M., Kojadinovic, I., Maechler, M., Yan, J. (2015), copula: Multivariate Dependence with Copulas. R package version 0.999-14. http://CRAN.R-project.org/package=copula

Hofert, M., Maechler, M. (2011), Nested Archimedean Copulas Meet R: The nacopula Package, *J. Stat. Softw. 39*(9), 1-20. http://www.jstatsoft.org/v39/i09/.

IPCC (2013), Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge Univ.Press, Cambridge.

Joe, H. (1997), Multivariate models and dependence concept. Chapman and Hall, New York.

Kao, S.-C., Govindaraju, R.S. (2010), A copula-based joint deficit index for droughts, *J. Hydrol. 380*(1), 121-134.

Kojadinovic, I., Yan, J. (2010), Modeling Multivariate Distributions with Continuous Margins Using the copula R Package, *J. Stat. Softw. 34*(9), 1-20. http://www.jstatsoft.org/v34/i09/.

Lee, J., Nadolnyak, D. (2012), The impacts of climate change on agricultural farm profits in the U.S. Paper presented at the Agricultural & Applied Economics Association's annual meeting, Seattle.

Lee, T., Modarres, R., Ouarda, T. (2013), Data-based analysis of bivariate copula tail dependence for drought duration and severity, *Hydrol. Process. 27*(10), 1454-1463.

Li, L., Li, W., Barros, A. (2013), Atmospheric moisture budget and its regulation of the summer precipitation variability over the Southeastern United States, *Clim. Dynam. 41*, 613–631.

Li, W., Li, L., Fu, R., Deng, Y., Wang, H. (2010), Changes to the North Atlantic subtropical high and its role in the intensification of summer rainfall variability in the Southeastern United States, *J. Clim. 24*, 1499–1506.

Li, Y., Ye, W., Wang, M., Yan, X. (2009), Climate change and drought: A risk assessment of crop-yield impacts, *Clim. Res. 39*(1), 31-46.

Liu, M., Xu, X., Sun, A.Y., Wang, K., Liu, W., Zhang, X. (2014), Is southwestern China experiencing more frequent precipitation extremes? *Environ. Res. Lett. 9*(6), 064002.

Liu, X., Wang, S., Zhou, Y., Wang, F., Yang, G., Liu, W. (2016), Spatial analysis of meteorological drought return periods in China using Copulas, *Nat. Hazards 80*(1), 367-388.

Ma, M., Ren, L., Singh, V.P., Tu, X., Jiang, S., Liu, Y. (2015), Evaluation and application of the SPDI-JDI for droughts in Texas, USA, *J. Hydrol. 521*, 34-45.

Ma, M., Ren, L., Singh, V.P., Yang, X., Yuan, F., Jiang, S. (2014), New variants of the Palmer drought scheme capable of integrated utility, *J. Hydrol. 519*, 1108-1119.

MacDonald, G. M., et al. (2008), Climate Warming and 21st-Century Drought in Southwestern North America, *Eos Trans. AGU 89*(9), 82–82, doi:10.1029/2008EO090003.

Madadgar, S., Moradkhani, H. (2013), Drought analysis under climate change using copula, *J. Hydrol. Eng. 18*(7), 746-759.

Marsaglia, G., Tsang, W.W., Wang, J. (2003), Evaluating Kolmogorov's distribution, *J. Stat. Softw. 8*(18). http://www.jstatsoft.org/v08/i18/.

Masud, M., Khaliq, M., Wheater, H. (2015), Analysis of meteorological droughts for the Saskatchewan River Basin using univariate and bivariate approaches, *J. Hydrol. 522*, 452-466.

Madhu, S., Kumar, T.P.L., Barbosa, H., Rao, K.K., Bhaskar, V. (2015), Trend analysis of evapotranspiration and its response to droughts over India, *Theor. Appl. Climatol. 121*, 41-51.

McCabe, G. J., Palecki, M. A., Betancourt, J. L. (2004), Pacific and Atlantic Ocean influences on multidecadal drought frequency in the United States, *Proc. Natl. Acad. Sci. USA 101*, 4136–4141.

Mishra, A. K., Desai, V. R. (2005), Spatial and temporal drought analysis in the Kansabati River Basin. India, *Int. J. River Basin Manage 3*(1), 31-41.

Mishra, A. K., Singh, V. P. (2008), Development of drought SAF curves. In: Singh, V.P. (Ed.), Hydrology and Hydraulics, Water Resour. Publ., Highlands Ranch, Colo, pp. 811–831.

Mishra, A.K., Singh, V.P. (2010), Drought modeling-A review, *J. Hydrol. 391*(1-2), 202-216.

Obasi, G.O.P. (1994), WMO's role in the international decade for natural disaster reduction, *Bull. Am. Meteorol. Soc. 75*(9), 1655-1661.

Palmer, W. (1965), Meteorological Drought. Research paper no.45, U.S. Department of Commerce Weather Bureau, February 1965 (58 pgs). Available online by the NOAANational Climatic Data Center at http://www.ncdc.noaa.gov/temp-and-precip/drought/docs/palmer.pdf

Quiring, S.M., Papakyriakou, T.N. (2003), An evaluation of agricultural drought indices for the Canadian prairies, *Agr. Forest Meteorol. 118*(1-2), 49–62.

Rajsekhar, D., Singh, V.P., Mishra, A.K. (2015), Multivariate drought index: An information theory based approach for integrated drought assessment, *J. Hydrol. 526*, 164-182.

Reddy, M.J., Ganguli, P. (2012), Application of copulas for derivation of drought severity–duration–frequency curves, *Hydrol. Process. 26*, 1672-1685.

Sadri, S., Burn, D.H. (2014), Copula-Based Pooled Frequency Analysis of Droughts in the Canadian Prairies, *J. Hydrol. Eng. 19*(2), 277-289.

Saghafian, B., Mehdikhani, H. (2014), Drought characterization using a new copula-based trivariate approach, *Nat. Hazards 72*(3), 1391-1407.

Schubert, S. D., Suarez, M. J., Pegion, P. J., Koster, R. D., Bacmeister, J. T. (2004), Causes of Long-Term Drought in the U.S. Great Plains, *J. Clim. 17*(3), 485-503.

Serinaldi, F., Grimaldi, S. (2007), Fully nested 3-copula: Procedure and application on hydrological data, *J. Hydrol. Eng. 12*(4), 420-430.

Sheffield, J., Andreadis, K. M., Wood, E. F., Lettenmaier, D. P. (2009), Global and Continental Drought in the Second Half of the Twentieth Century: Severity–Area–Duration Analysis and Temporal Variability of Large-Scale Events, *J. Clim. 22*, 1962–1981.

Sheffield, J., Wood, E.F., Roderick, M.L. (2012), Little change in global drought over the past 60 years, *Nature 491*(7424), 435.

Shiau, J.T. (2003), Return period of bivariate distributed extreme hydrological events, *Stoch. Env. Res. Risk 17*(1-2), 42-57.

Shiau, J.T. (2006), Fitting drought duration and severity with two-dimensional copulas, *Water Resour. Manag. 20*(5), 795-815.

Shiau, J.T., Shen, H.W. (2001), Recurrence analysis of hydrologic droughts of differing severity, *J. Water Resour. Plann. Manag. 127*(1), 30–40.

Sklar, A. (1959), Fonctions de répartitions à n dimensions et leurs marges, *Publ. Inst. Stat. Univ. Paris 8*, 229–231.

Soule, P.T. (1992), Spatial patterns of drought frequency and duration in the contiguous USA based on multiple drought event definitions, *Int. J. Climatol. 12*, 11-24.

Trenberth, K.E., Dai, A., van der Schrier, G., Jones, P.D., Barichivich, J., Briffa, K.R., Sheffield, J. (2013), Global warming and changes in drought, *Nat. Clim. Change 4*(1), 17-22.

van der Schrier, G., Barichivich, J., Briffa, K.R., Jones, P.D. (2013), A scPDSI-based global data set of dry and wet spells for 1901–2009, *J. Geophys. Res. Atmos. 118*, 4025–4048.

Vicente-Serrano, S. M. (2006). Differences in spatial patterns of drought on different time scales: an analysis of the Iberian Peninsula, *Water Resour. Manage. 20*, 37-60.

Vicente-Serrano, S. M., Beguería, S., López-Moreno, J. I. (2010), A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index, *J. Clim. 23*, 1696–1718.

Wang, H., Fu, R., Kumar, A., Li, W. (2010), Intensification of Summer Rainfall Variability in the Southeastern United States during Recent Decades, *J. Hydrometeorol.11*(4), 1007-1018.

Weiss, J. L., Castro, C. L., Overpeck, J. T. (2009), Distinguishing Pronounced Droughts in the Southwestern United States: Seasonality and Effects of Warmer Temperatures, *J. Clim. 22*, 5918–5932.

Wells, N., Goddard, S., Hayes, M.J. (2004), A self-calibrating Palmer Drought Severity Index, *J. Climate 17*(12), 2335-2351.

WMO report (2006), Drought monitoring and early warning: concepts, progress and future challenges. WMO report no. 1006, World Meteorological Organization.

Wong, G., Lambert, M.F., Leonard, M., Metcalfe, A.V. (2010), Drought Analysis Using Trivariate Copulas Conditional on Climatic States, *J. Hydrol. Eng. 15*(2), 129-141.

Xu, K., Yang, D., Xu, X., Lei, H. (2015), Copula based drought frequency analysis considering the spatio-temporal variability in Southwest China, *J. Hydrol. 527*, 630-640.

Yan, H., Wang, S., Wang, J., Lu, H., Guo, A., Zhu, A., Myneni, R., Shugart, H.H. (2016), Assessing spatiotemporal variation of drought in China and its impact on agriculture

during 1982–2011 by using PDSI indices and agriculture drought survey data, *J. Geophys. Res. Atmos. 121,* D024285.

Yan, J. (2007), Enjoy the Joy of Copulas: With a Package copula, *J. Stat. Softw. 21*(4), 1-21. http://www.jstatsoft.org/v21/i04/.

Zhang, L., Singh, V.P. (2007), Gumbel-Hougaard copula for trivariate rainfall frequency analysis, *J. Hydrol. Eng. 12,* 409-419.

Chapter 5 Conclusions

This thesis develops methods to detect changes in climate (i.e., non-stationarity) and their impact on crop yields. Especially drought as an extreme climate event has been examined in terms of its spatial and temporal change patterns using northern India and the Continental U.S. (CONUS) as case studies. First, the statistical method that uses the combined Pettitt test and the Bayesian inference detection method successfully identifies the types of changes (i.e., abrupt or gradual) and the significant change point(s) between two linear segments that have different slopes, intercepts or variances. The mixed trends and abrupt shifts identified in climate provide additional information beyond the usual assumption of linear trends. Applying the method to the time series of 1910-2009 of the CONUS, it is found that abrupt changes occurred more frequently with precipitation than with temperature. Hot spots of identified changes in climate show closely correlated spatial-temporal patterns. The identified spatial-seasonal variation and correlation of changes highlights the uncertainty and information loss of averaging climate variables over years and the assumption of linear trends. Moreover, ongoing trends are found with increasing temperature (except in summer in the Midwest), decreasing diurnal temperature range (DTR) (except in spring), and increasing precipitation in the Midwest (except in fall). Gradual trends and abrupt shifts require different adaptation measures, and making the distinction between the types of changes will help us make appropriate decisions in climate change adaptation, as well as understanding the causal mechanisms in climatic changes. In particular, the ongoing trends identified from historical records may imply continuous changes in the coming years, which may inform climate change mitigation and water and land management adaptations in the near future.

Furthermore, to assess the impact of changes in climate on crop yields, historical records during 1947-2010 in Nebraska, United States are used to assess the impact of changes in climate on irrigated and rainfed maize yield. The assessment is based on statistical methods that detect both gradual and abrupt changes in climate time series and determine their impacts on irrigated and rainfed crop yields separately. It is shown that a downward parabolic relationship of crop yield with the various climate variables including the mean diurnal temperature range (DTR) during a crop growth stage exists for both irrigated and rainfed maize, and the threshold points of the parabolic curves are different for irrigated and rainfed crops. In most Nebraska counties, a linearly decreasing trend of mean daily maximum and average temperature benefited both irrigated and rainfed maize during June and July, but with more benefits seen for the rainfed maize. The changes in mean summertime precipitation during the past half century caused a decline in the irrigated yield but an increase in the rainfed yield, which is likely related to excessive wetness due to more frequent occurrences of situations in which irrigation applications under dryness are quickly followed by wetness. In general, we find that the changes in mean climate variables during the past half century benefited less or hurt more the irrigated maize than the rainfed maize in Nebraska. That is to say, irrigation may aggravate the negative impacts of climate change. Although average irrigated maize yield was overall higher than the rainfed yield during 1947-2010, weather vagaries could considerably diminish the expected return of irrigation to crops.

The next part of this thesis particularly examines drought, an extreme climate event in terms its spatial and temporal patterns of multiple characteristics (duration, severity and intensity). Following the debate on whether drought has become more severe under climate change, this paper assesses drought frequency in northern and eastern India using two datasets of

Palmer Drought Severity Index (PDSI) (generated by Dai, 2013 and Sheffield et al., 2012). The univariate return period for three drought characteristics (duration, severity and peak intensity) is examined regarding whether drought has occurred with longer duration, higher severity and/or larger peak intensity. The spatial variation of those changes is analysed through eight areas in the study region. The temporal and spatial comparisons based on the univariate return period show different change patterns of duration, severity and peak intensity in different areas. Generally, in the areas which plant wheat more than rice, drought has been alleviated in duration and intensity after 1955; while in the areas which plant more rice than wheat, drought have been aggravated in duration, severity and intensity (except for area 8, a coastal area). This spatial change pattern may imply potential crop pattern change, for example, switching from rice to wheat in areas 3-7. Furthermore, the bivariate return period for pairs of drought characteristics based on the copulas and considering correlation between the drought characteristics is examined to understand how bivariate return periods change over time and space. Finally, it is also found that one data set (Sheffield et al.) results in more severe, longer and more intense drought in most of the areas, especially for the drought events with long-return-periods than the other (Dai).

Another case study of drought study is to investigate the spatial and temporal patterns of multiple drought characteristics (duration, severity and intensity) under different return periods during 1900 to 2012 in CONUS. We find two significant patterns: Pattern I shows persistent droughts in Western, Eastern U.S. and the Great Plains, which experienced large variations in the drought characteristics over long time; Pattern II shows transient droughts in the interior of CONUS, which experienced short-term variations in drought characteristics. The classification of change patterns of drought reveals the areas with persistent drought conditions (Pattern I), which may cause more damage to human and natural systems. In addition, spatial distributions of

duration, severity and intensity for short-term and long-term return periods in the same time period are found to be different. Trend analysis shows that duration, severity and intensity of droughts under the various return periods are increasing in most of the Pattern I regions. Moreover, trends in these drought characteristics at long and short return periods are different at some locations, showing the different trends of extreme and mild droughts.

In future, more work could be conducted by following the work of this thesis. First, how climate change or technology advancement has affected the variance of crop yield can be addressed. In the thesis, the trend of crop yield due to technology advancement or soil types have been analyzed; also, the effects of climate change on crop yield have been analyzed. However the impact assessment assumes that the relationship between climate variables and crop yields remain unchanged from the past to the future. This assumption may not hold, for example, technology advancement can increase the resistance of crop to adverse weather (e.g., drought), and the relationship between climate variables and crop yields could vary with time. Taking the 2012 drought as example, which significantly decreased crop yield, if the same drought event happened in earlier years or in the future, whether it would affect yields more or less is a question. The hypothesis of variance change would affect institutions, for example the crop insurance model which can be improved by a probability distribution of crop yields and then the distribution of crop losses.

In addition, the analysis of spatial and temporal patterns of a climate variable (e.g., drought) can be extended to a hydrological variable, e.g., the return period of streamflow. This extension will further answer the question of how non-stationary climate and hydrology will affect water resources engineering design and management. In a stationary climate, classic concepts of return period assume that in design, by just specifying year T, one can compute the

probability of exceeding u_T in a time period of length N, for any N (i.e., the distribution is the same through all years) (Rootzen and Katz, 2013). Under the assumption of stationarity, the hydrologic risk management only needs to specify a standard used for all years in terms of a high quantile corresponding to the specified probability of failure for a single year (Institute for Water Resources, 2011). The concepts served the engineering community well in designing in the past. However, due to the non-stationarity in climate, these concepts based on the assumption of stationarity may not be applicable (Milly et al., 2008). Katz (2010) argued that shifts in extremes can be more reliably derived indirectly from changes in the overall probability distribution of a climate variable (e.g., shifts in the mean and standard deviation) than through direct statistical modeling of extremes. Therefore, methodology is needed to detect the change for abrupt shifts and gradual trends in parameters of probability distributions, with which the estimated probability distributions can be distinguished for each year breaking the assumption of classic concepts of return period. A risk of failure under non-stationarity over a time period much longer than a year (e.g., the probability of at least one extreme event over the lifetime of a structure) based on the fitted probability distribution for each year in the chosen period is more practical and useful. Therefore, it is worthy of addressing what the risk caused by non-stationarity to a design based on stationarity is; if the trend in parameters of probability distribution continues, what the near-future design level should be.

5.1 References

Institute for Water Resources (2011), Flood Risk Management Approaches: As Being Practiced in Japan, Netherlands, United Kingdom, and United States, IWR Report No. 2011-R-08, U.S. Army Corps of Engineers.

Katz, R. W. (2010), Statistics of extremes in climate change, *Clim. Change, 100*(1), 71-76.

Milly, P. C. D., J. Betancourt, M. Falkenmark, R. M. Hirsch, Z. W. Kundzewicz, D. P. Lettenmaier, and R. J. Stouffer (2008), Climate change - Stationarity is dead: Whither water management?, *Science*, *319*(5863), 573-574.

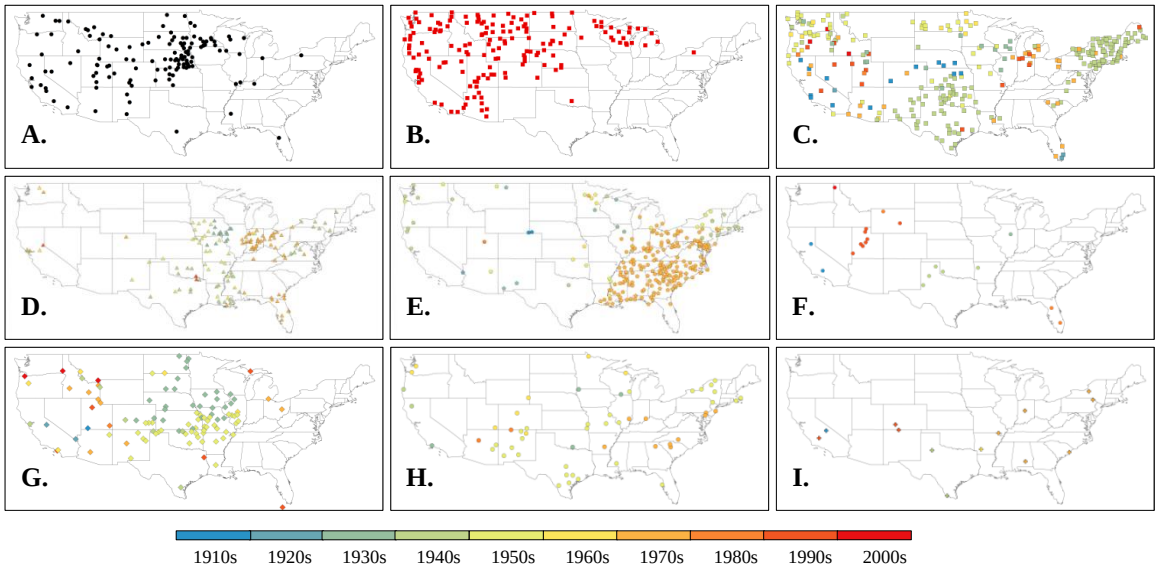
Rootzen, H., and R. W. Katz (2013), Design Life Level: Quantifying risk in a changing climate, *Water Resour. Res.*, *49*(9), 5964-5972.

Appendix A: Supporting Information for Chapter 2

Supporting Information includes:

Figures A.1-A.7

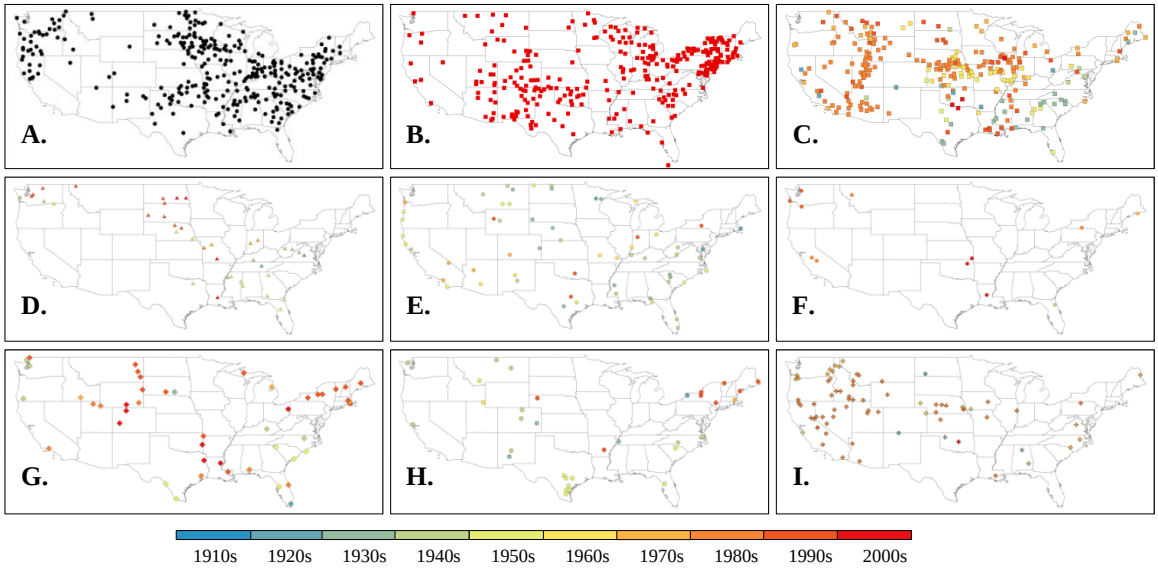
1) Winter mean maximum temperature



A. Type I.	B. Type II.	C. Type IV.I
D. Type IV.D	E. Type V.I/D or	F. Type VI.I/D or
G. Type VII.I/D or	H. Type VIII.I/D or	I. Type X.I/D or

Figure A.1 Spatial distribution of stations with nine major change patterns (as ordered by schematic table below) with seasonal mean maximum temperature during 1910-2009.

2) Spring mean maximum temperature



3) Summer mean maximum temperature

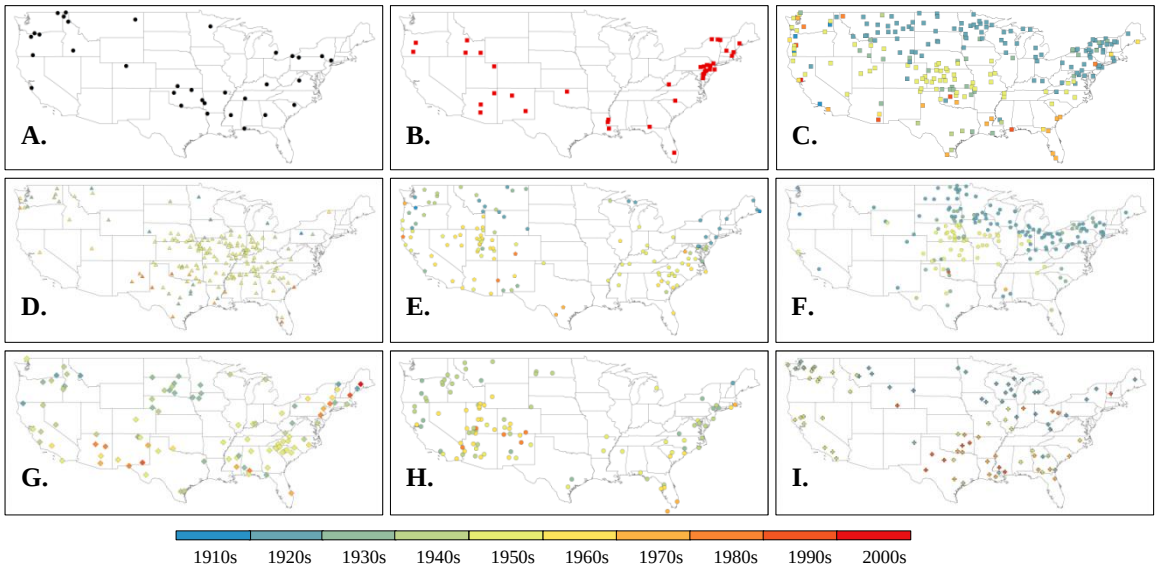


Figure A.1 (cont.)

4) Fall mean maximum temperature

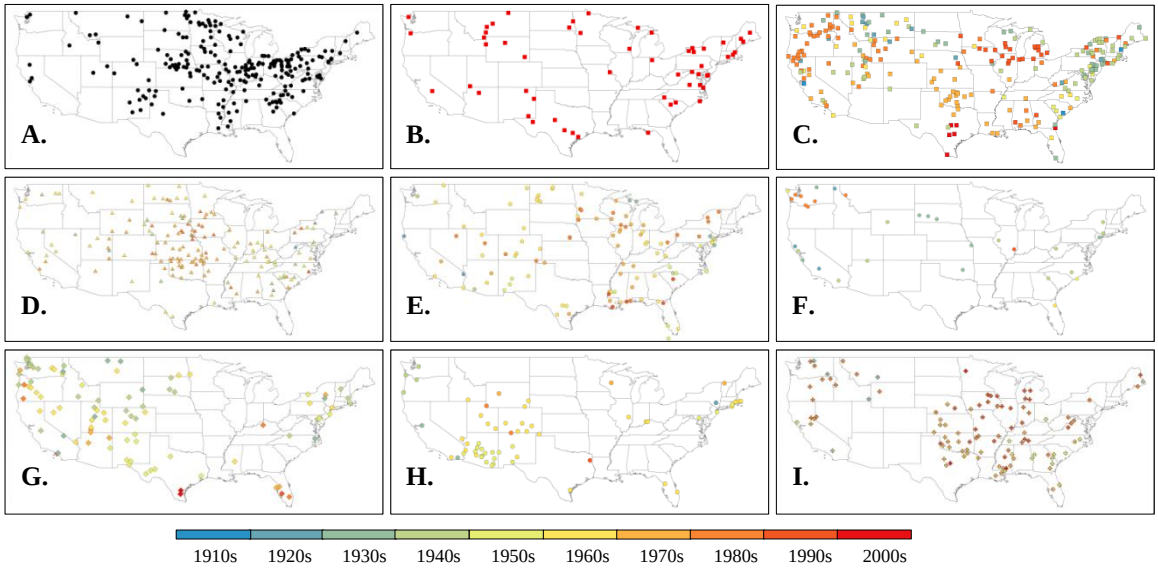
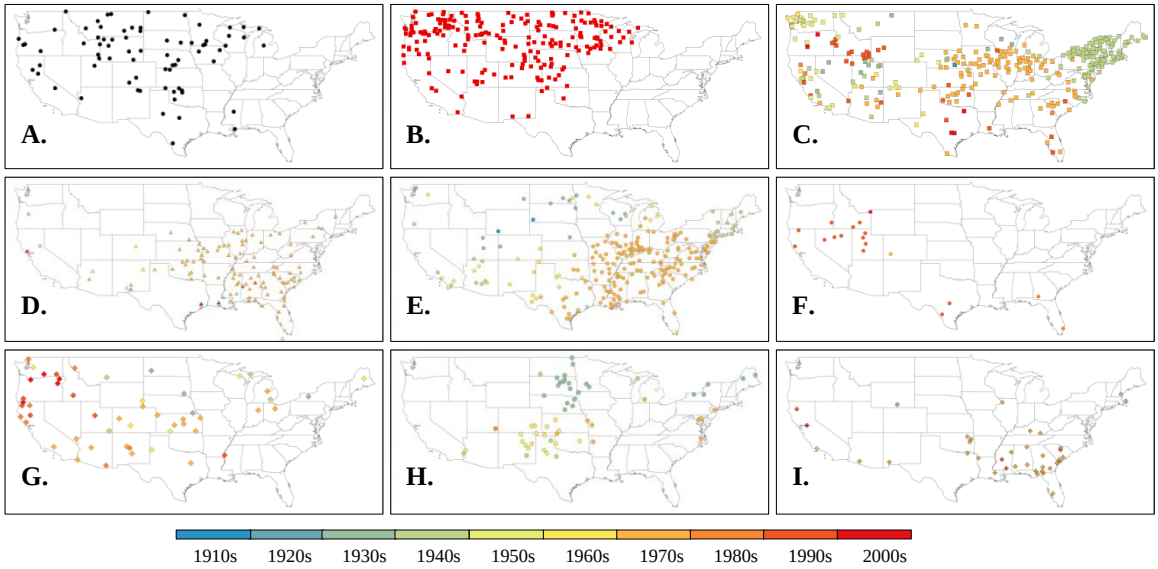
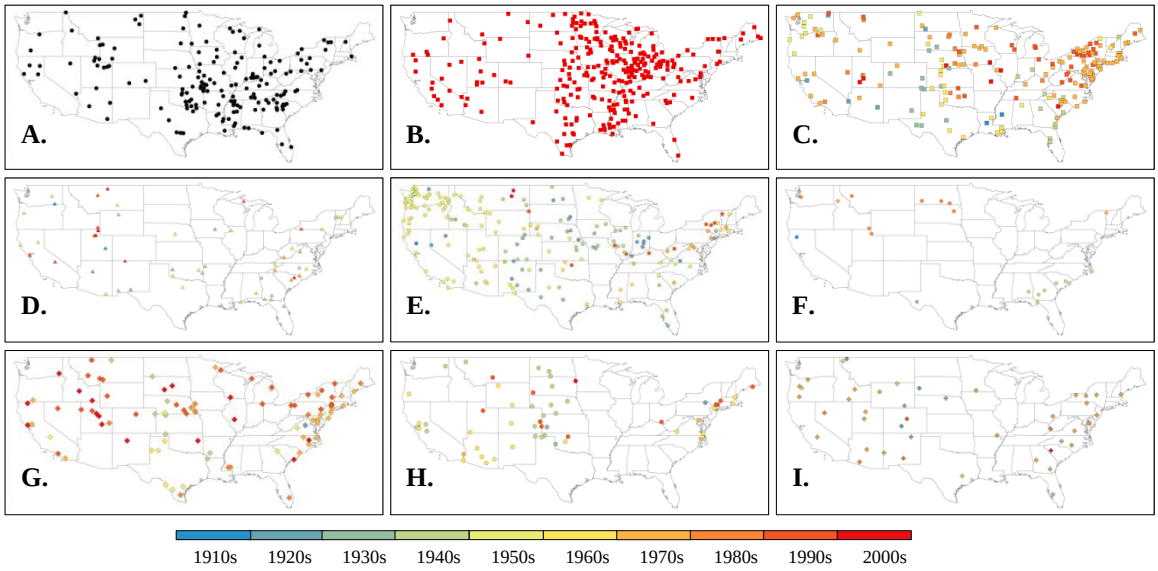


Figure A.1 (cont.)

1) Winter mean minimum temperature



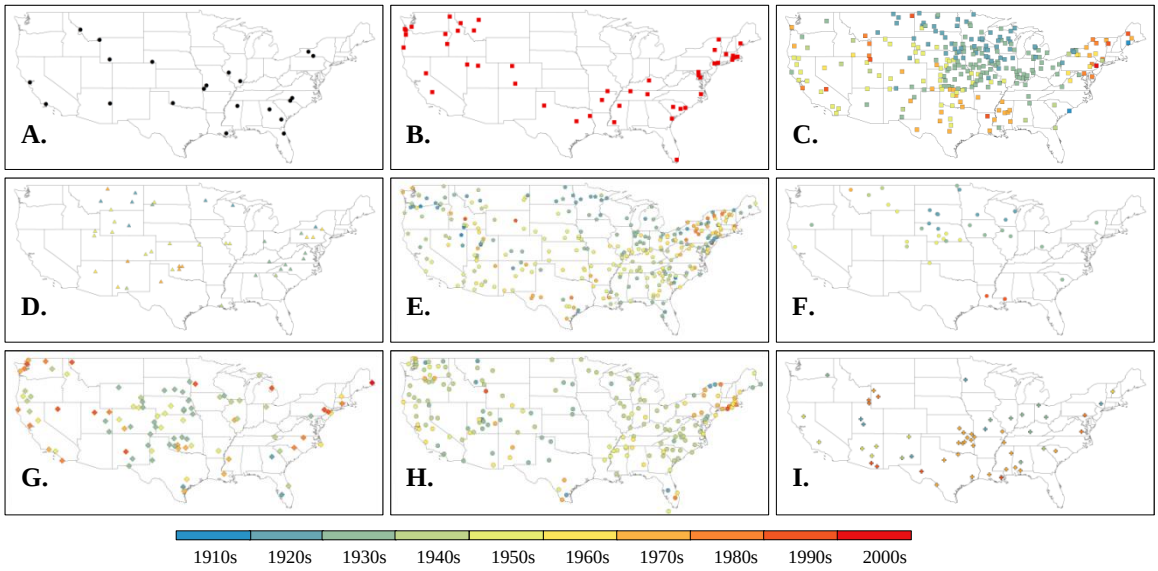
2) Spring mean minimum temperature



A. Type I.	B. Type II.	C. Type IV.I
D. Type IV.D	E. Type V.I/D or	F. Type VI.I/D or
G. Type VII.I/D or	H. Type VIII.I/D or	I. Type X.I/D or

Figure A.2 Spatial distribution of stations with nine major change patterns (as ordered by schematic table below) with seasonal mean minimum temperature during 1910-2009.

3) Summer mean minimum temperature



4) Fall mean minimum temperature

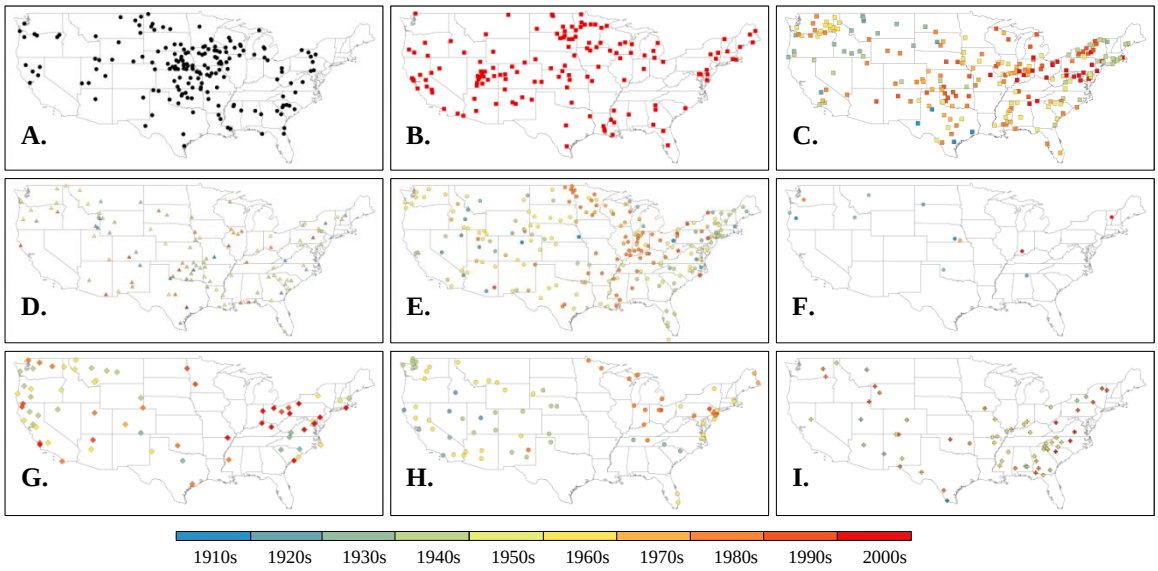
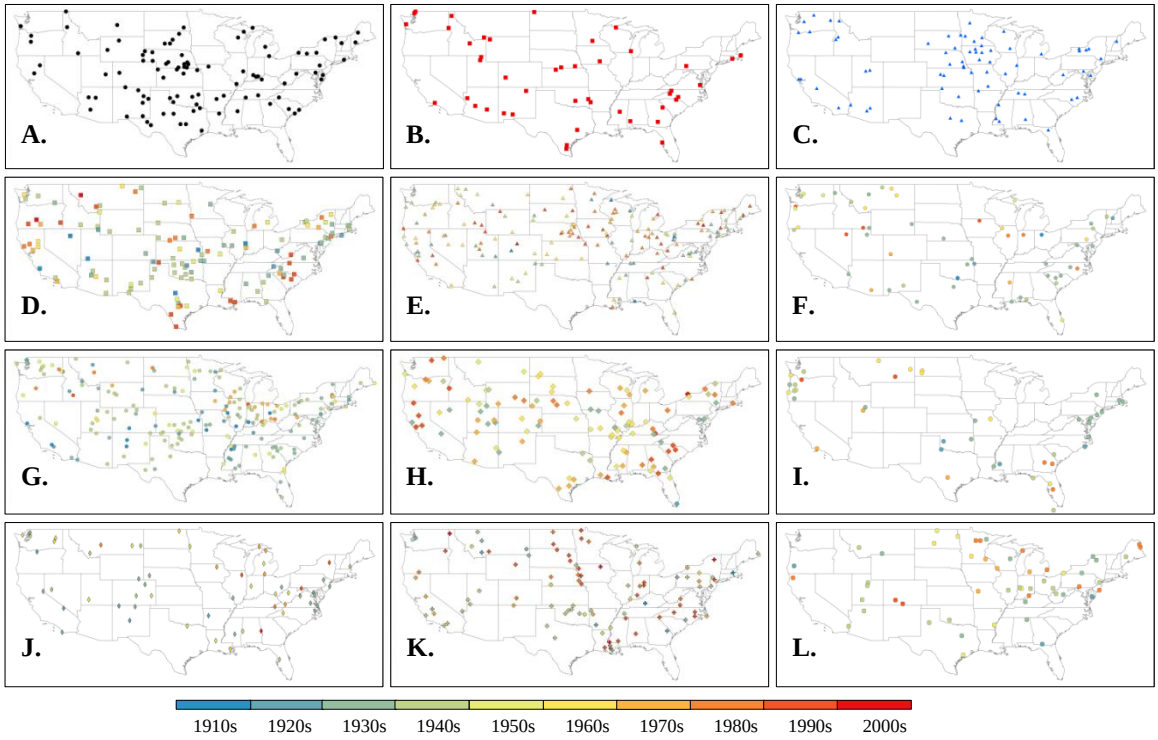


Figure A.2 (Cont.)

1) Winter mean DTR



A. Type I.	B. Type II.	C. Type III
D. Type IV.I	E. Type IV.D	F. Type V.I/D or
G. Type VI.I/D or	H. Type VII.I/D or	I. Type VIII.I/D or
J. Type IX.I/D or	K. Type X.I/D or	L. Type XI.I/D or

Figure A.3 Spatial distribution of stations with twelve major change patterns (as ordered by schematic table below) with seasonal mean DTR during 1910-2009.

2) Spring mean DTR

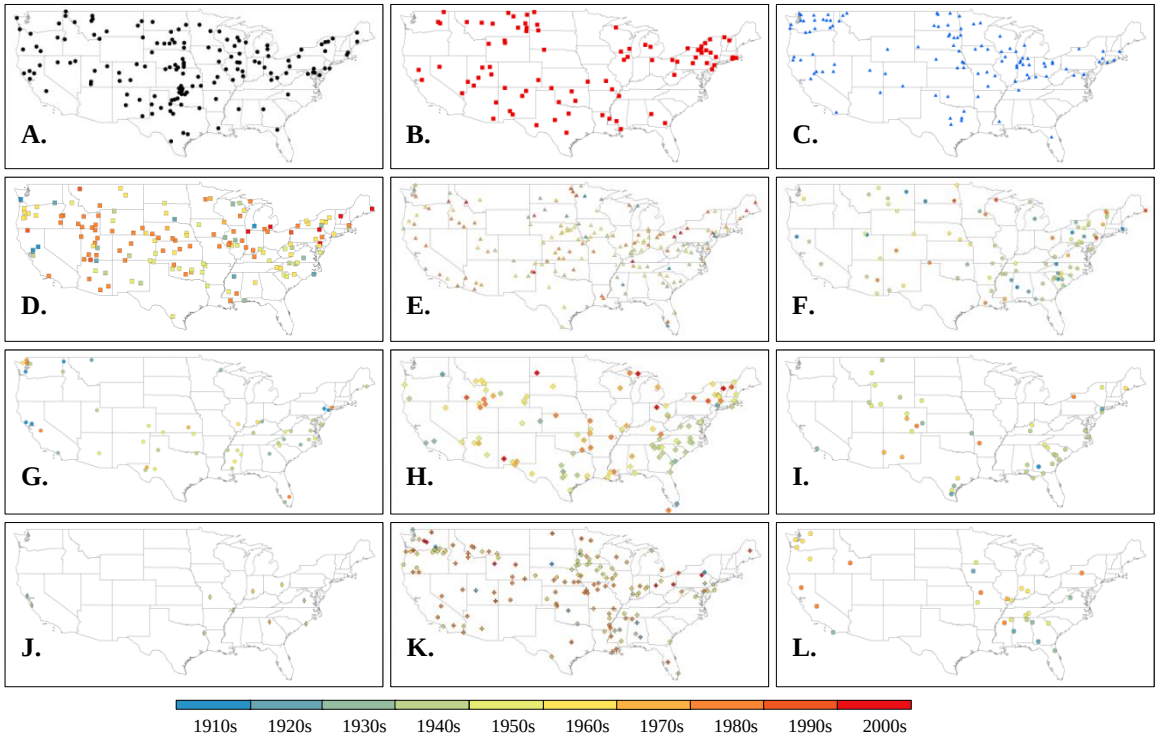


Figure A.3 (Cont.)

3) Summer mean DTR

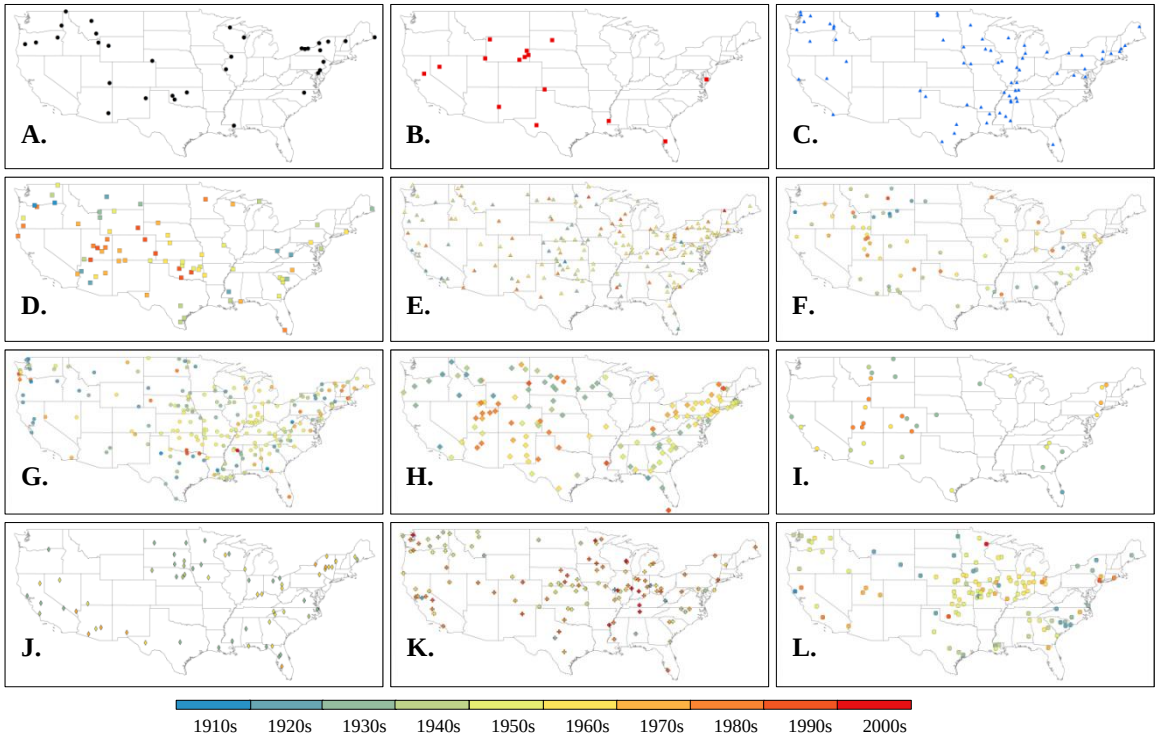


Figure A.3 (Cont.)

4) Fall mean DTR

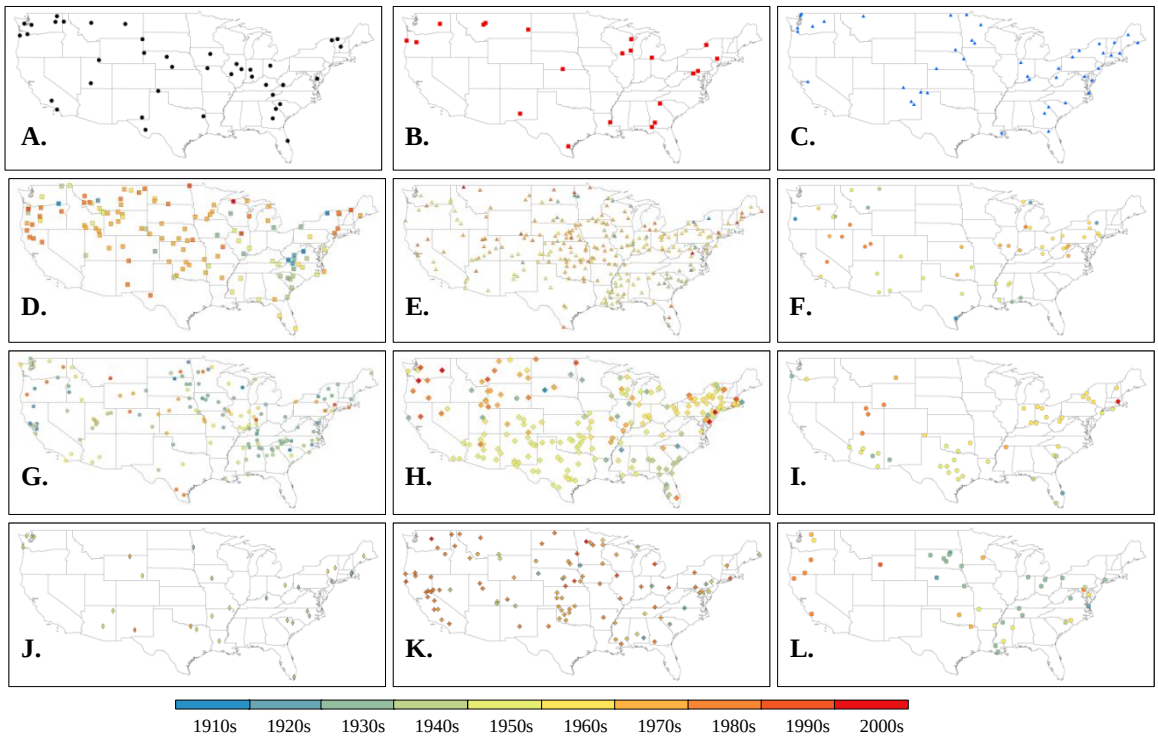
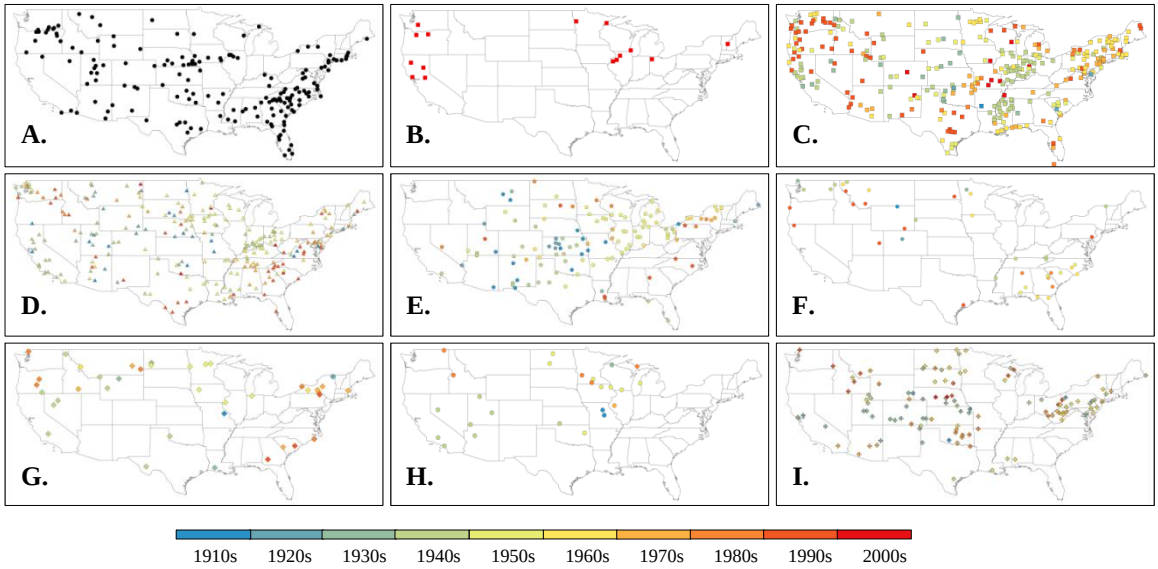
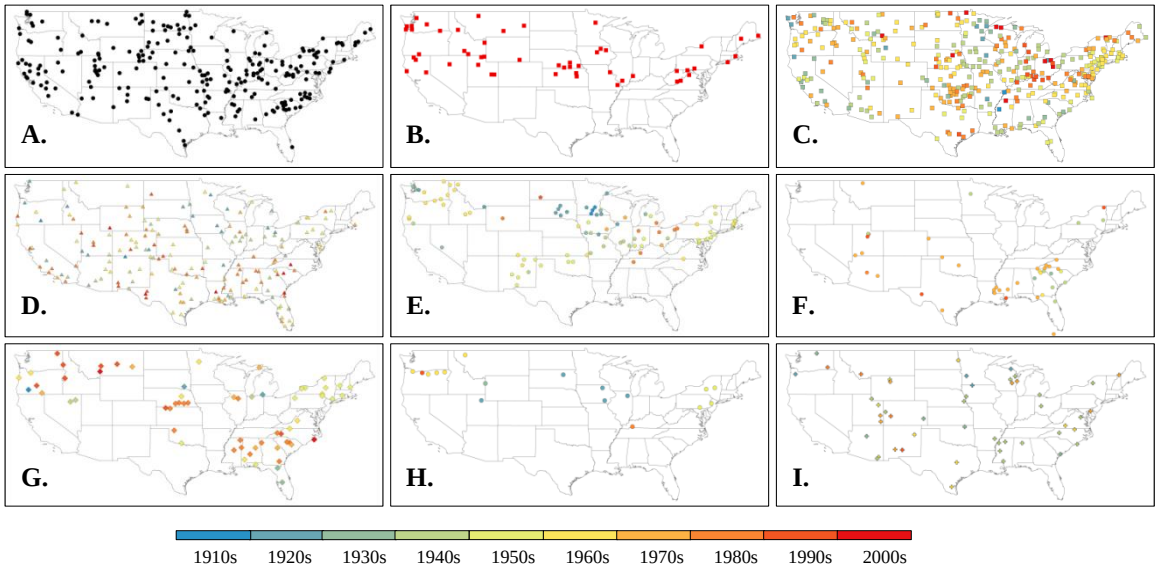


Figure A.3 (Cont.)

1) Winter total precipitation



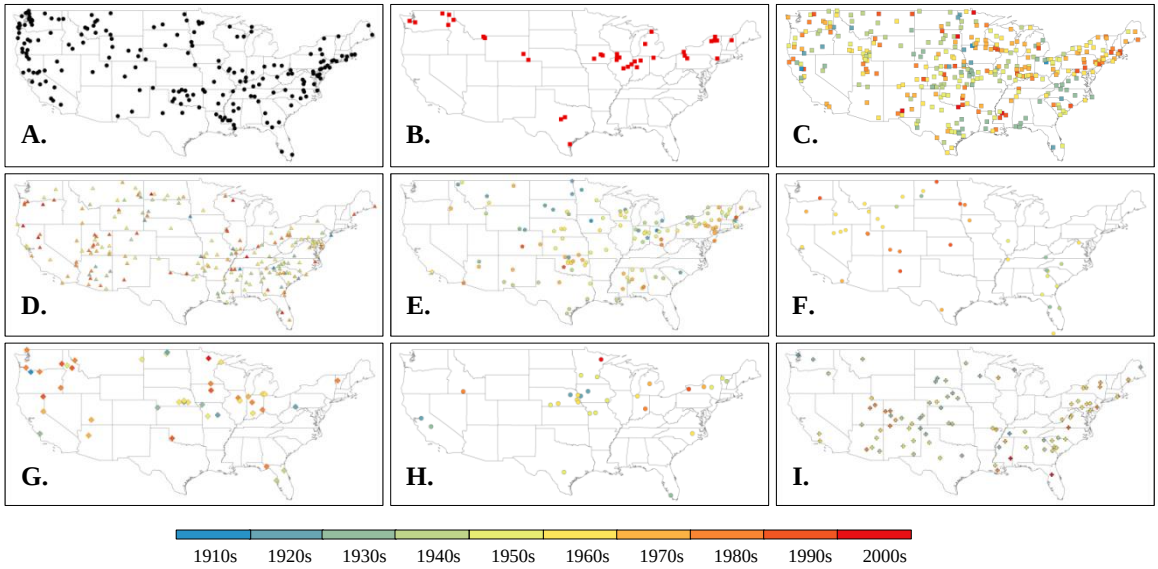
2) Spring total precipitation



A. Type I.	B. Type II.	C. Type IV.I
D. Type IV.D	E. Type V.I/D or	F. Type VI.I/D or
G. Type VII.I/D or	H. Type VIII.I/D or	I. Type X.I/D or

Figure A.4 Spatial distribution of stations with nine major change patterns (as ordered by schematic table below) with seasonal total precipitation during 1910-2009.

3) Summer total precipitation



4) Fall total precipitation

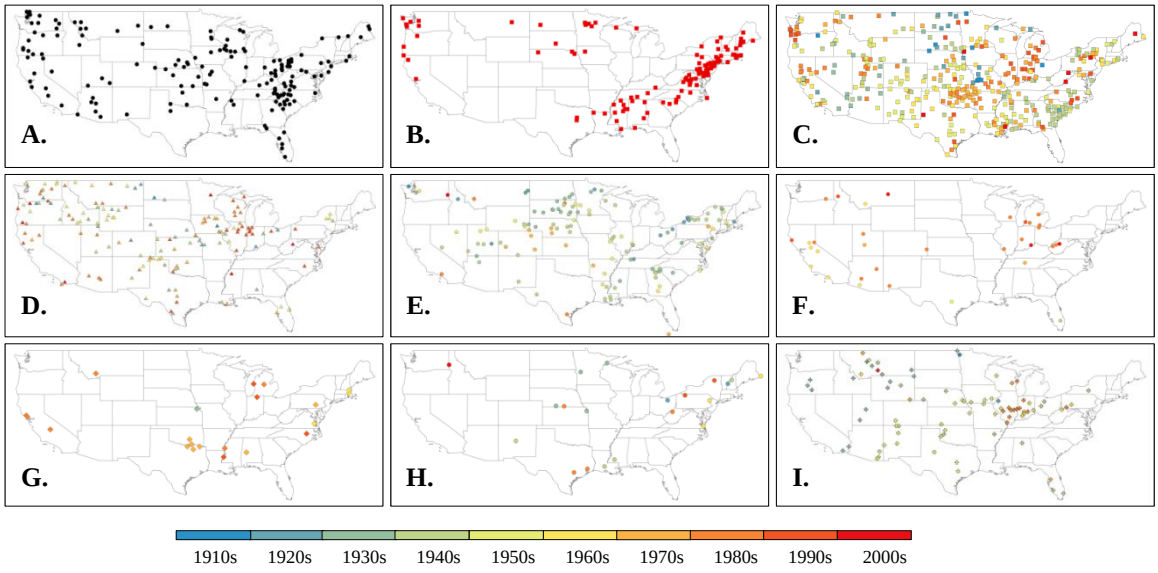
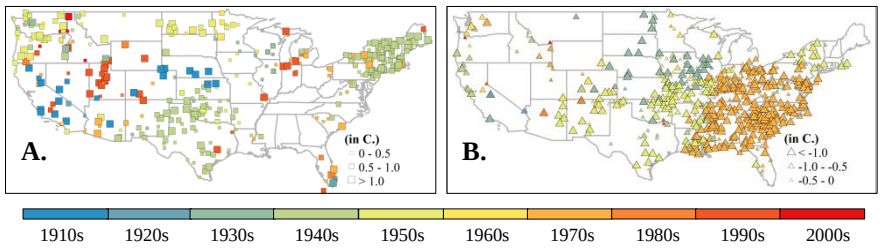
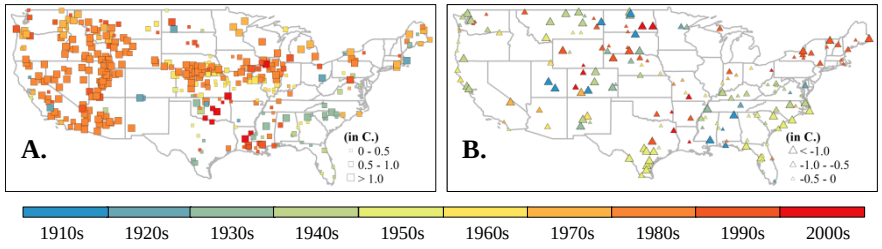


Figure A.4 (Cont.)

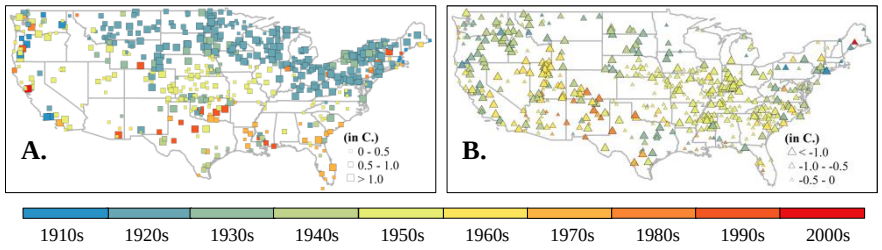
1) Winter mean maximum temperature



2) Spring mean maximum temperature



3) Summer mean maximum temperature



4) Fall mean maximum temperature

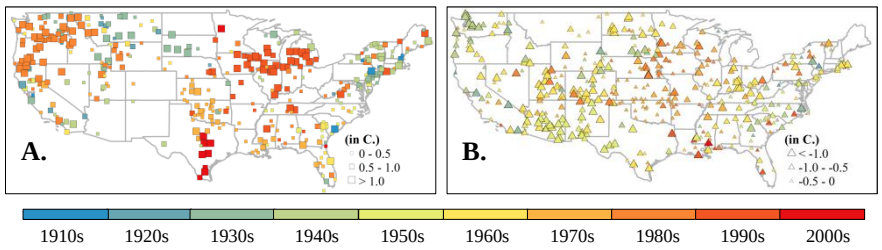
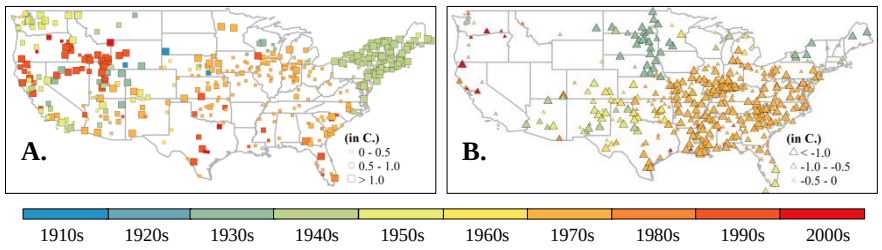
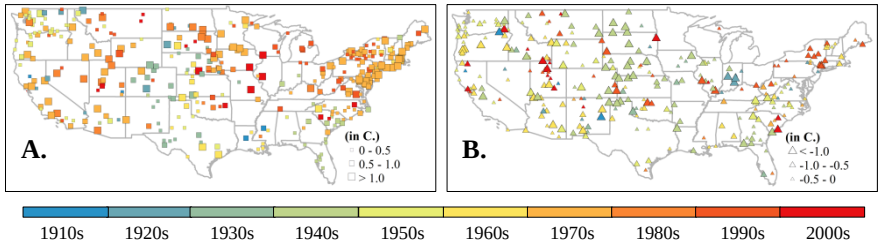


Figure A.5 Magnitude of (**A**) an abrupt upward change and (**B)** an abrupt downward change at an occurrence year (shown by colors) for a climatic variable during 1910-2009.

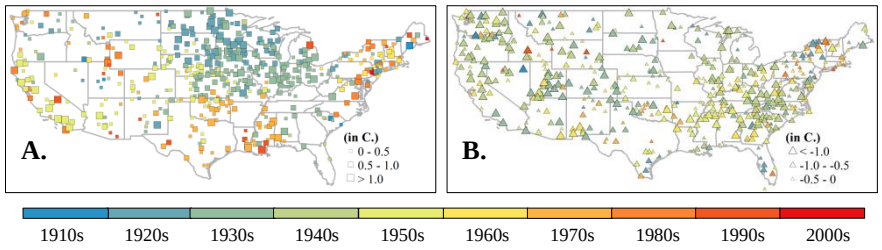
5) Winter mean minimum temperature



6) Spring mean minimum temperature



7) Summer mean minimum temperature



8) Fall mean minimum temperature

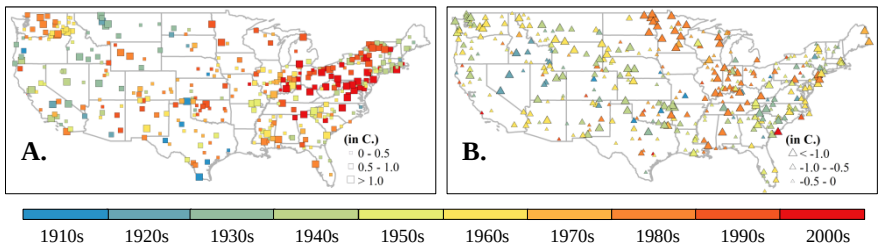
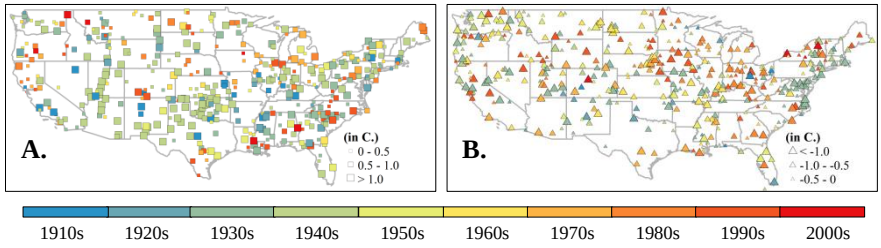
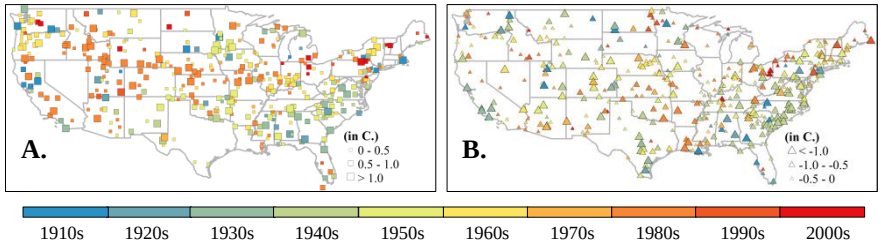


Figure A.5 (Cont.)

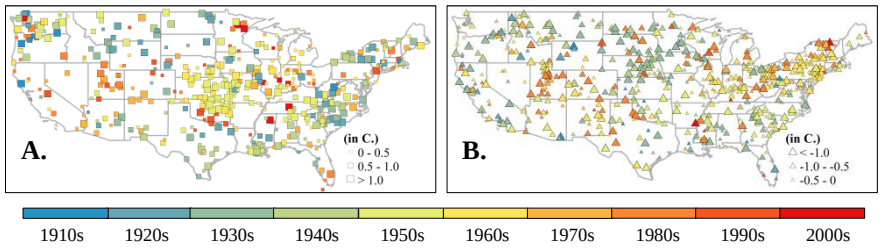
9) Winter mean DTR



10) Spring mean DTR



11) Summer mean DTR



12) Fall mean DTR

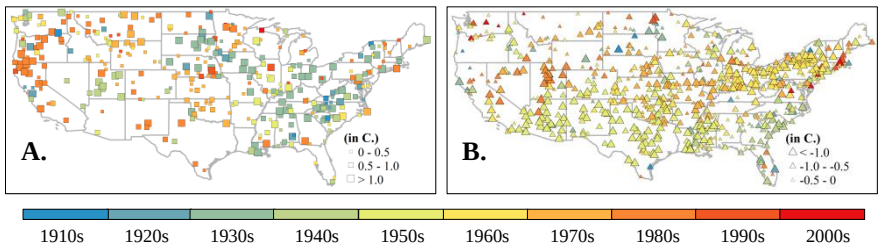
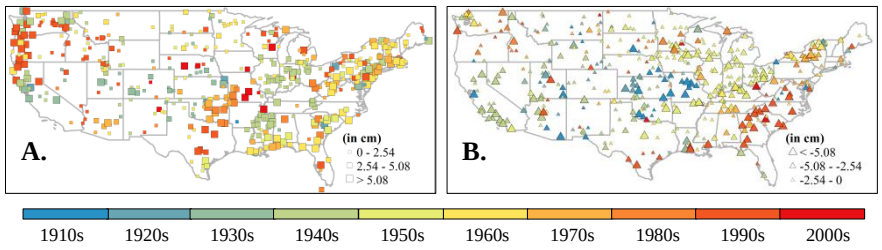
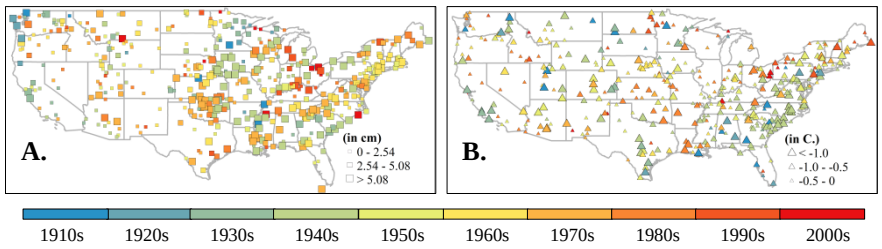


Figure A.5 (Cont.)

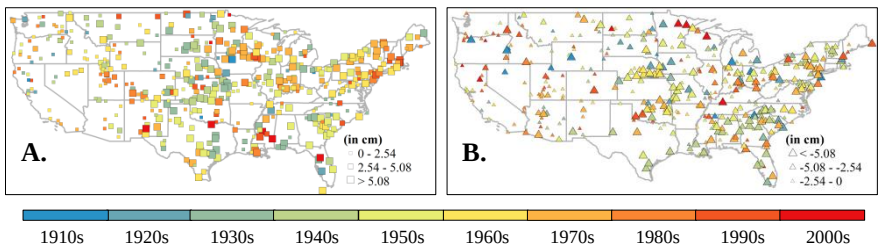
13) Winter total precipitation



14) Spring total precipitation



15) Summer total precipitation



16) Fall total precipitation

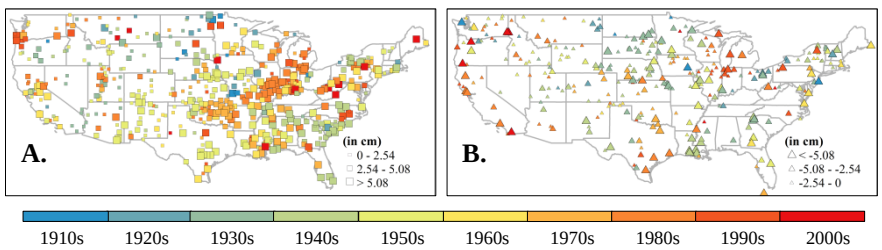


Figure A.5 (Cont.)

1) Winter mean maximum temperature

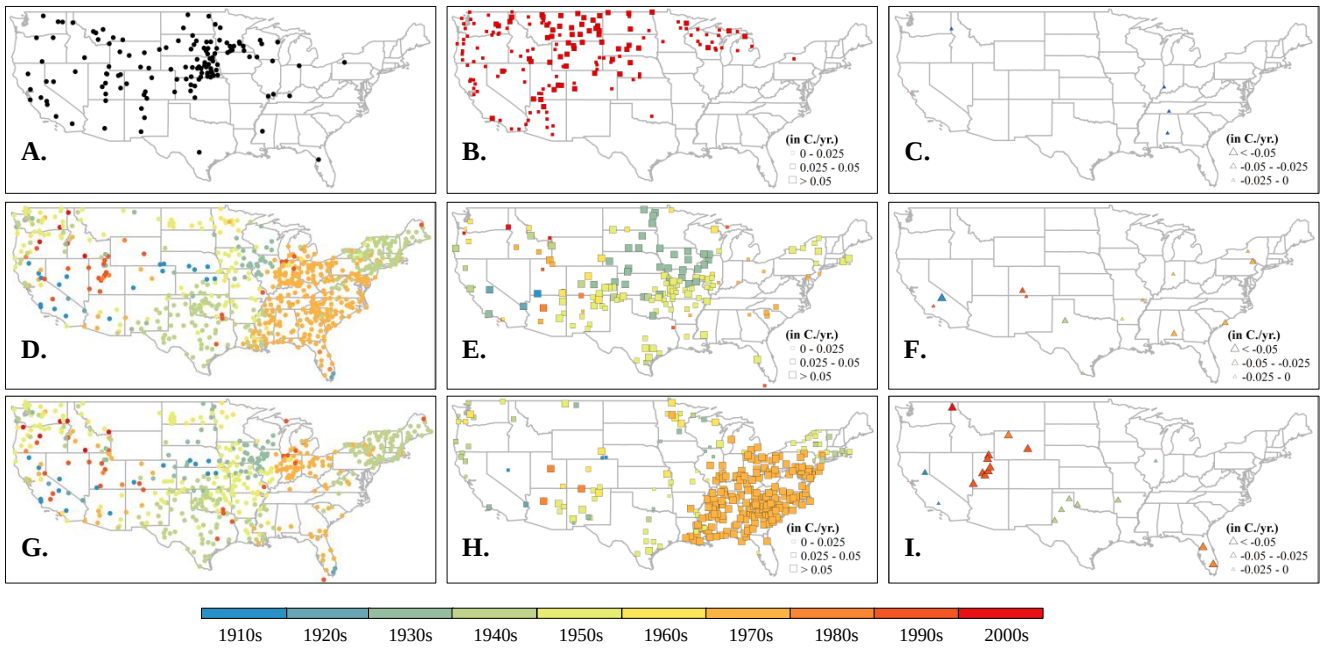
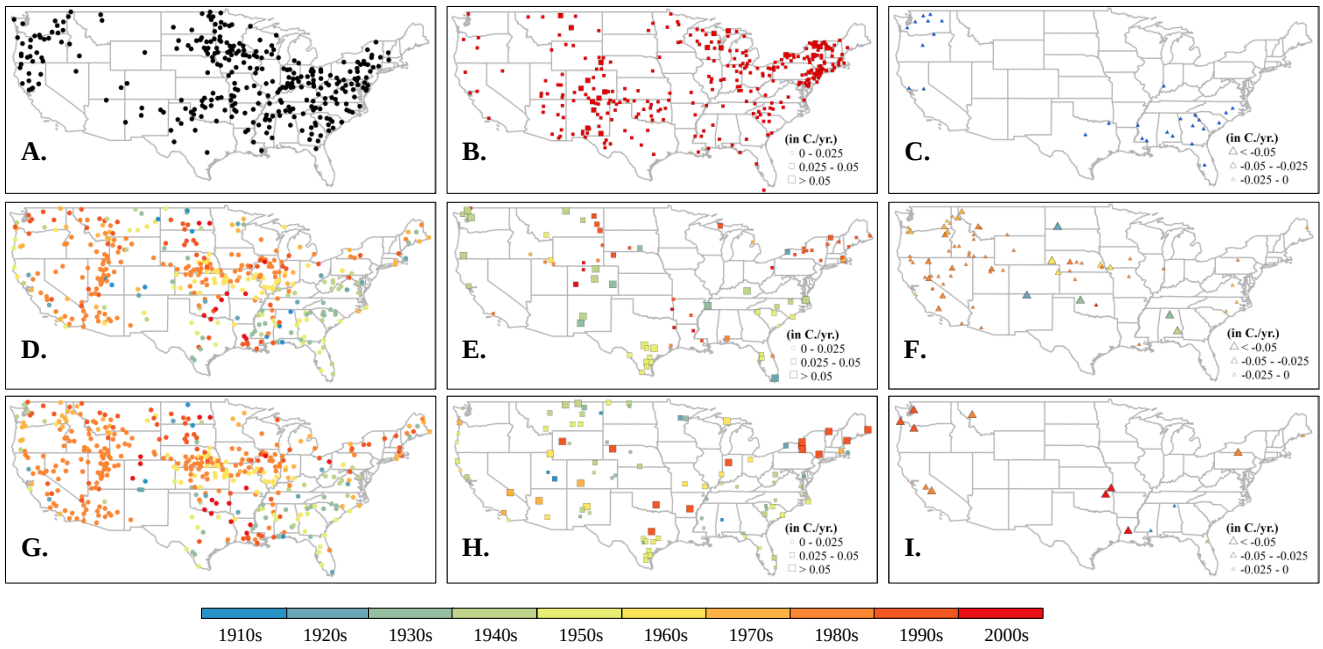


Figure A.6 Changes of a climatic variable during 1910-2009, (**A**) spatial distribution of "No change" in the entire period; (**B**) gradient of 100-year increasing trend; (**C**) gradient of 100-year decreasing trend; (**D**) spatial distribution of "No change" before an occurrence year; (**E**) gradient of a gradually increasing trend before an occurrence year; (**F**) gradient of a gradually decreasing trend before an occurrence year; (**G**) spatial distribution of "No change" after an occurrence year; (**H**) gradient of a gradually increasing trend after an occurrence year; (**I**) gradient of a gradually decreasing trend after an occurrence year.

2) Spring mean maximum temperature



3) Summer mean maximum temperature

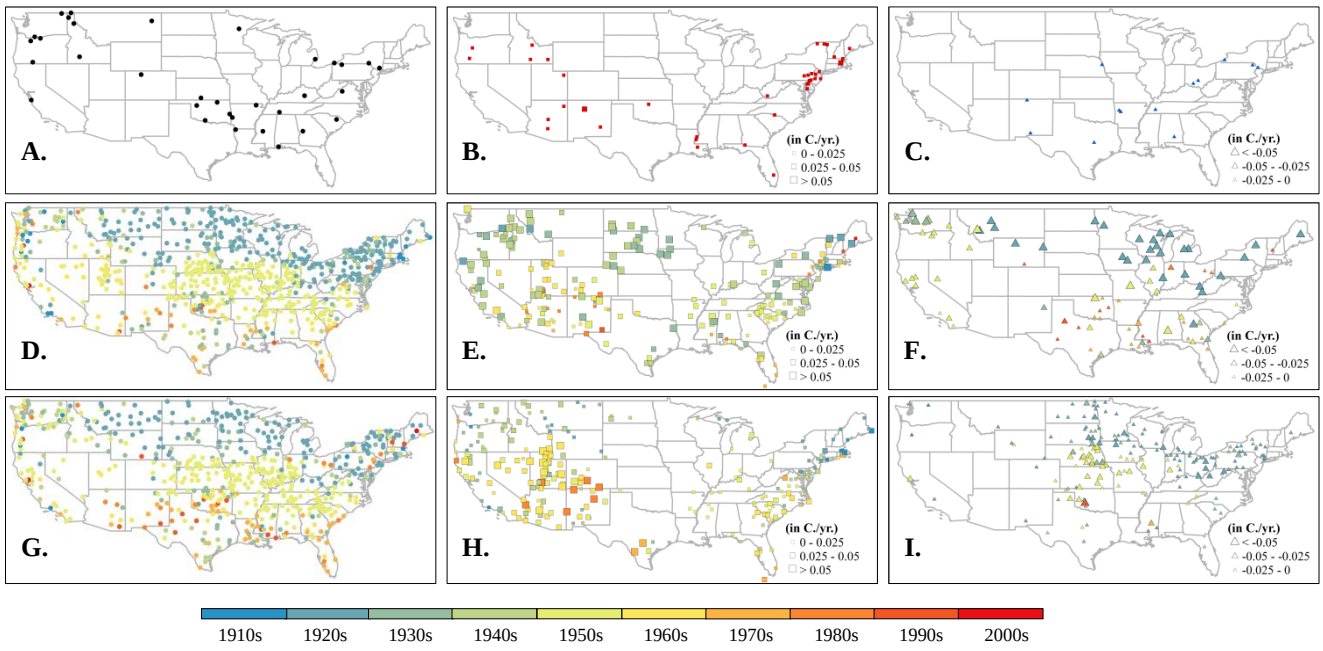
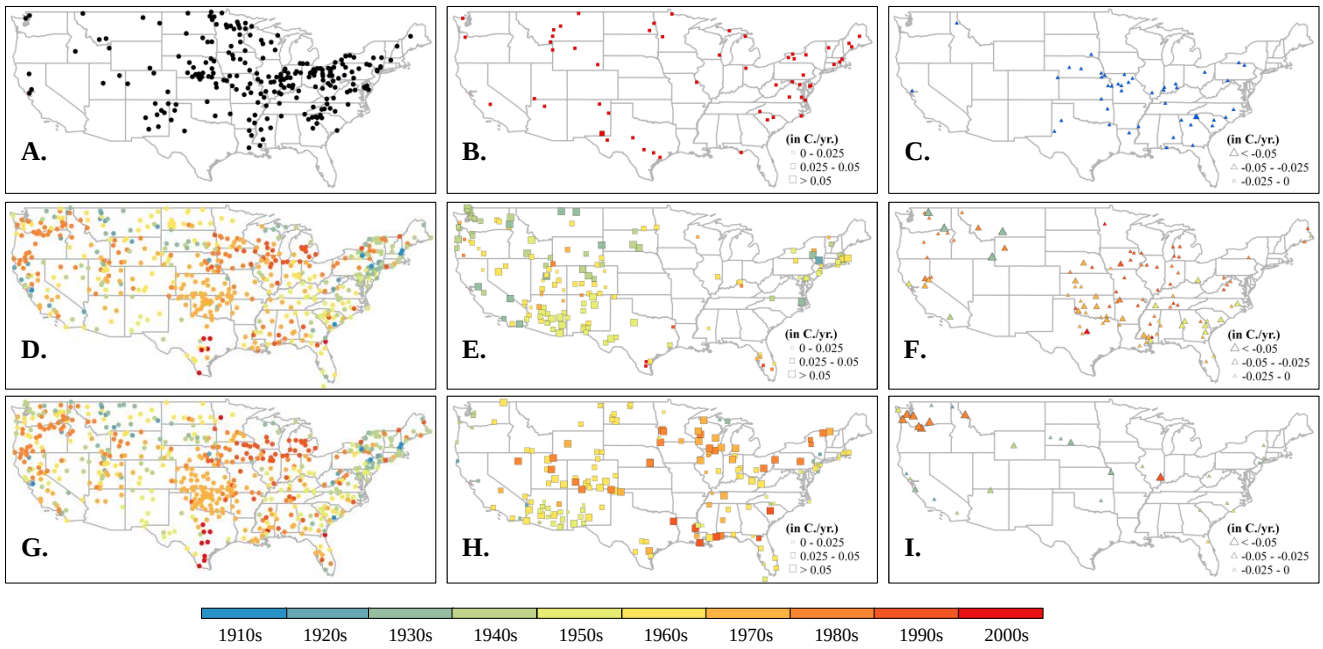


Figure A.6 (Cont.)

4) Fall mean maximum temperature



5) Winter mean minimum temperature

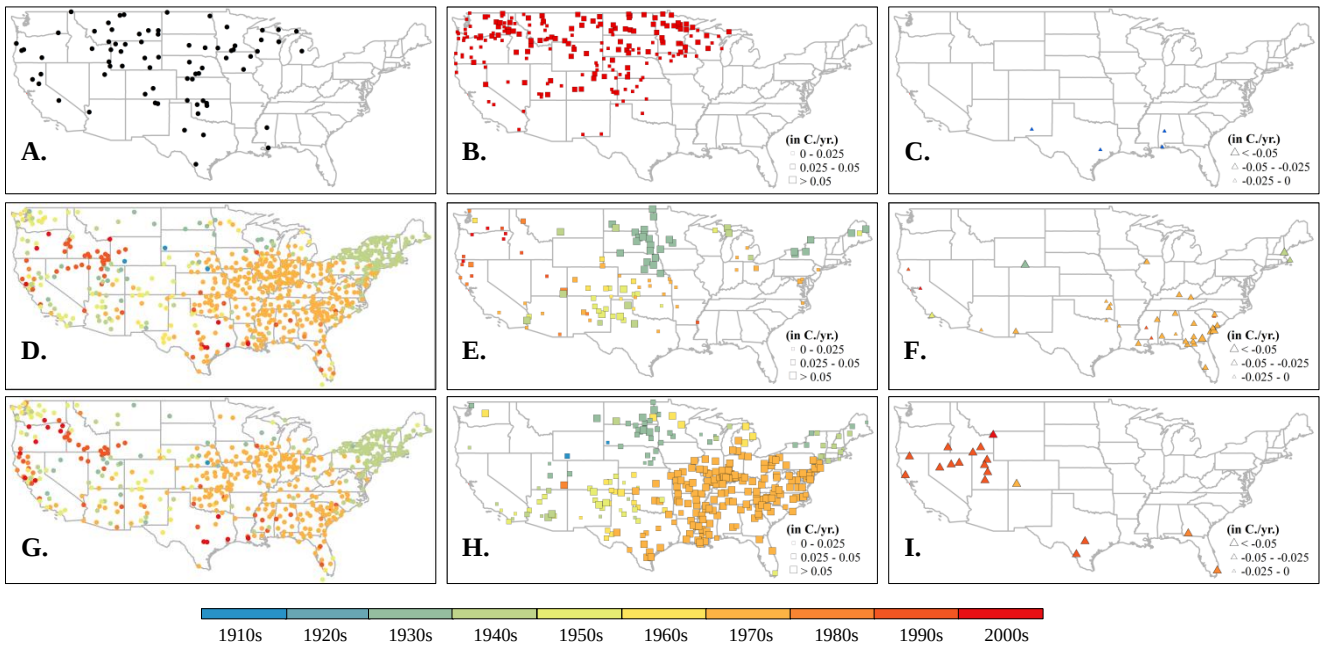
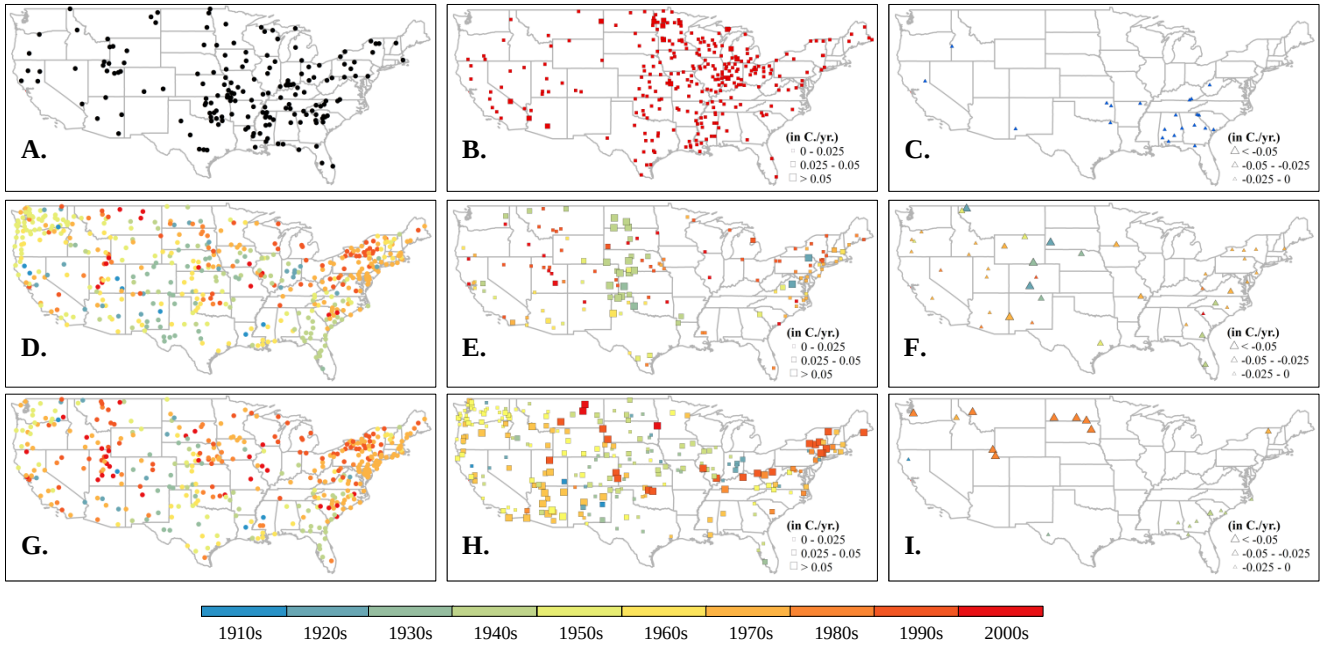


Figure A.6 (Cont.)

6) Spring mean minimum temperature



7) Summer mean minimum temperature

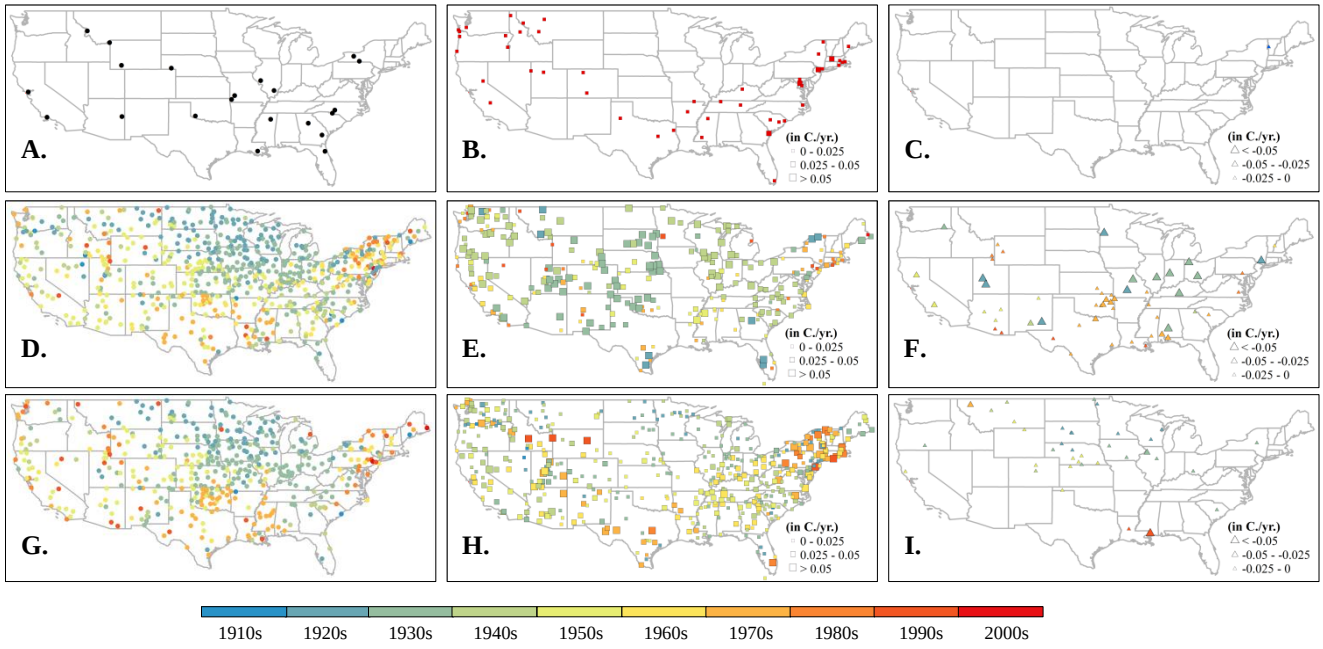
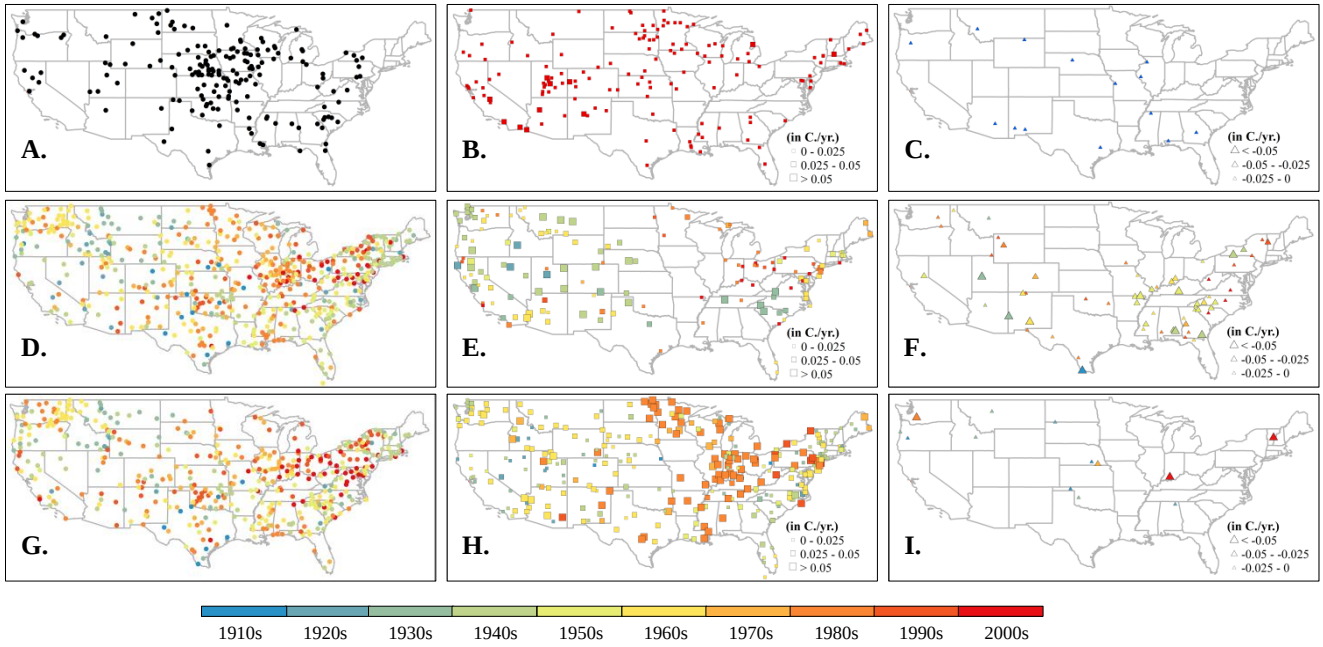


Figure A.6 (Cont.)

8) Fall mean minimum temperature



9) Winter mean DTR

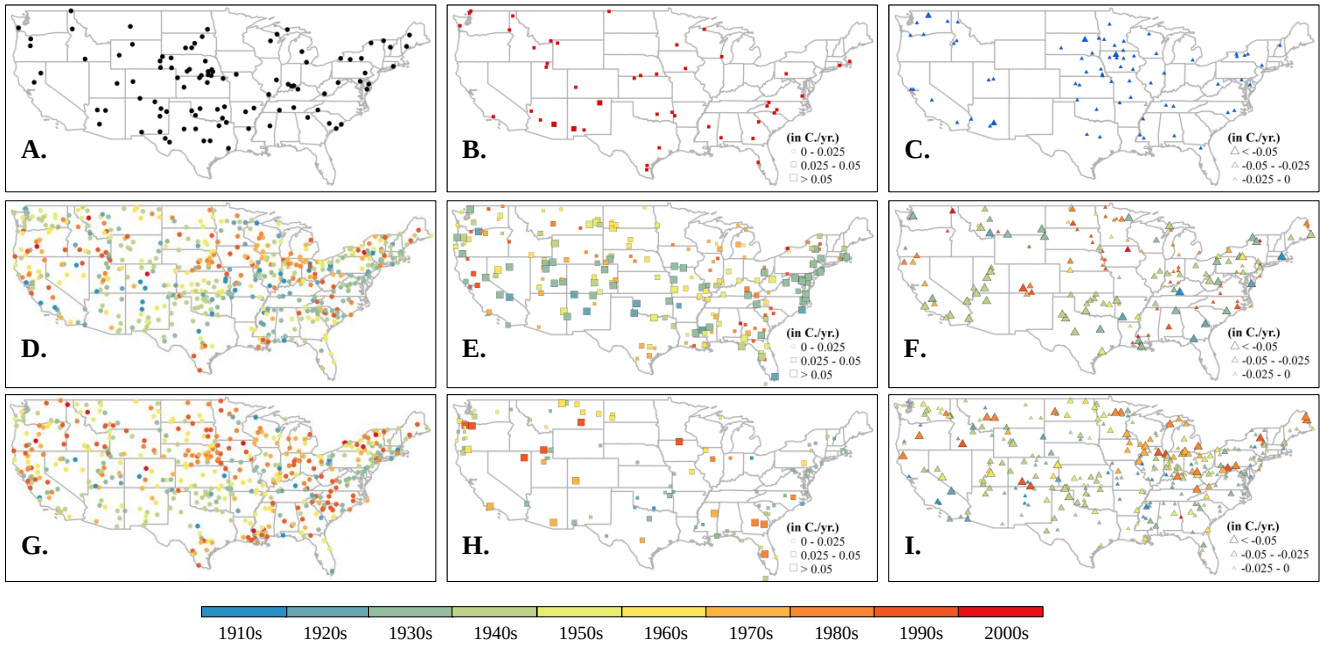
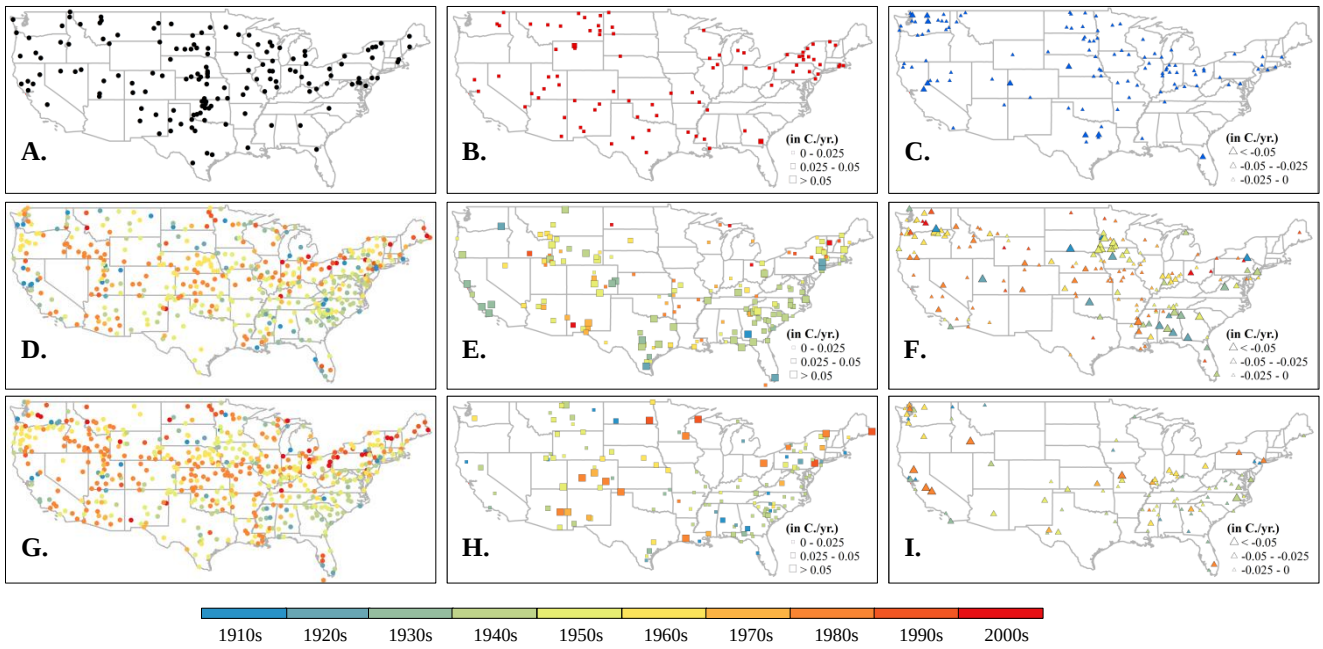


Figure A.6 (Cont.)

10) Spring mean DTR



11) Summer mean DTR

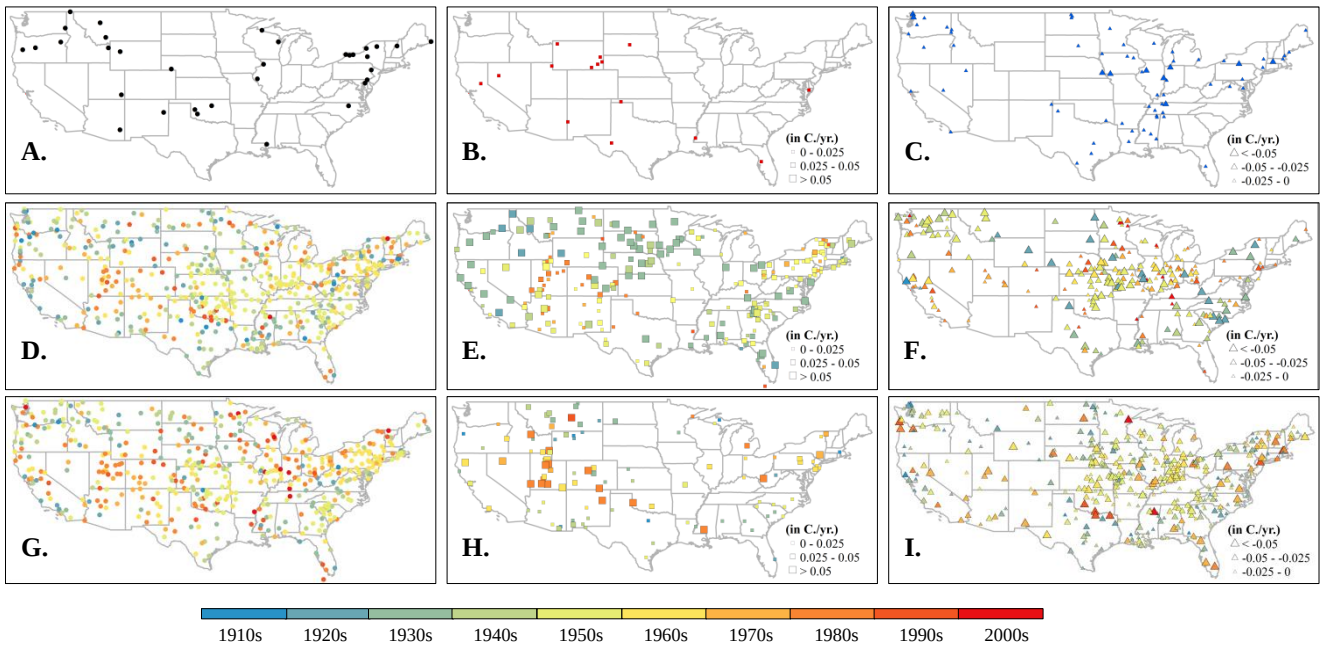
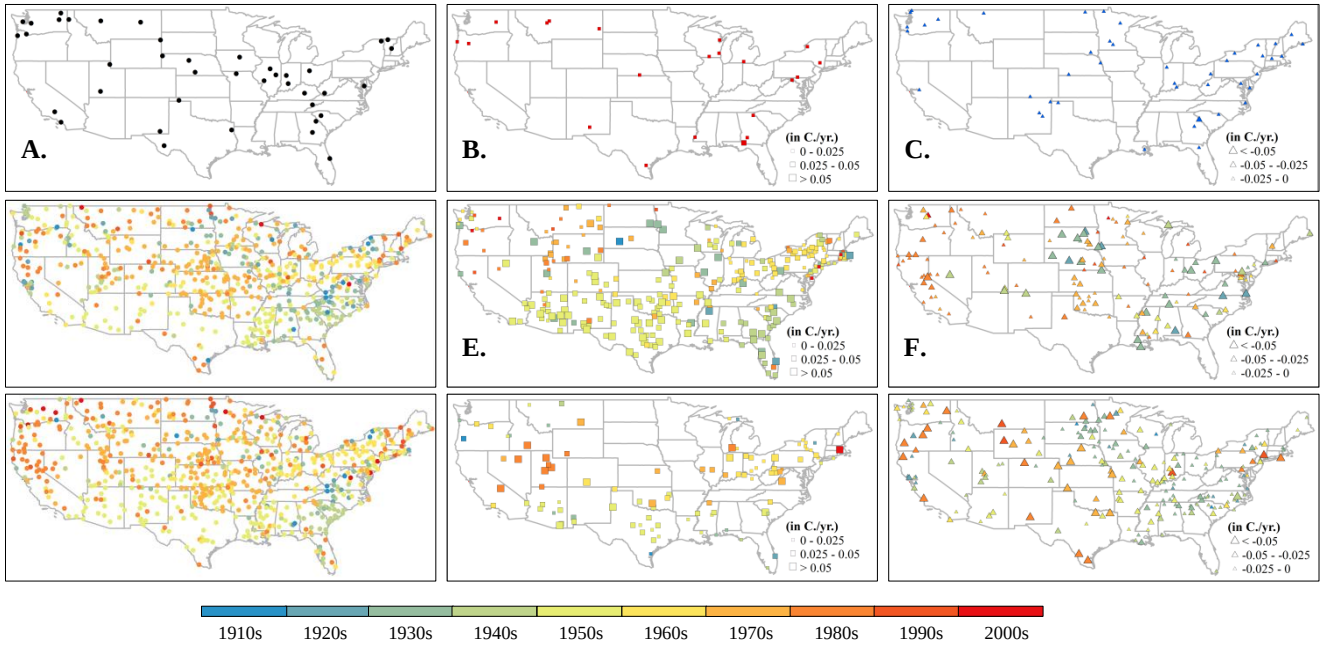


Figure A.6 (Cont.)

12) Fall mean DTR



13) Winter total precipitation

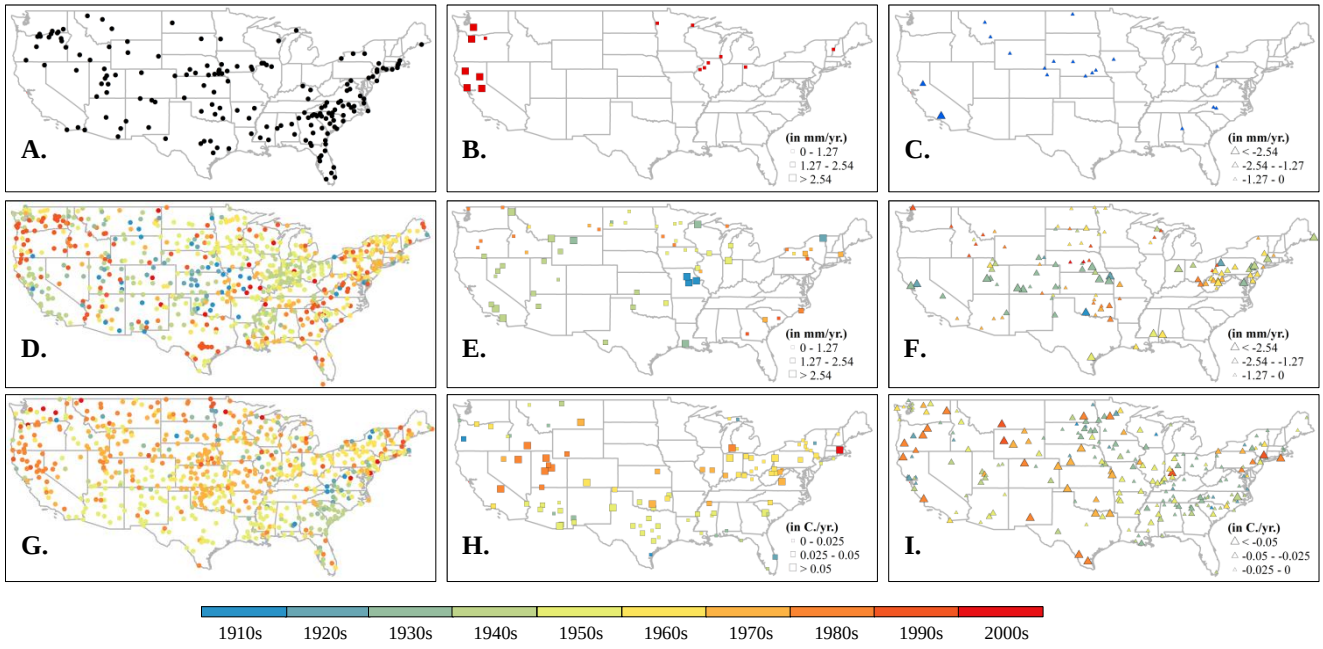
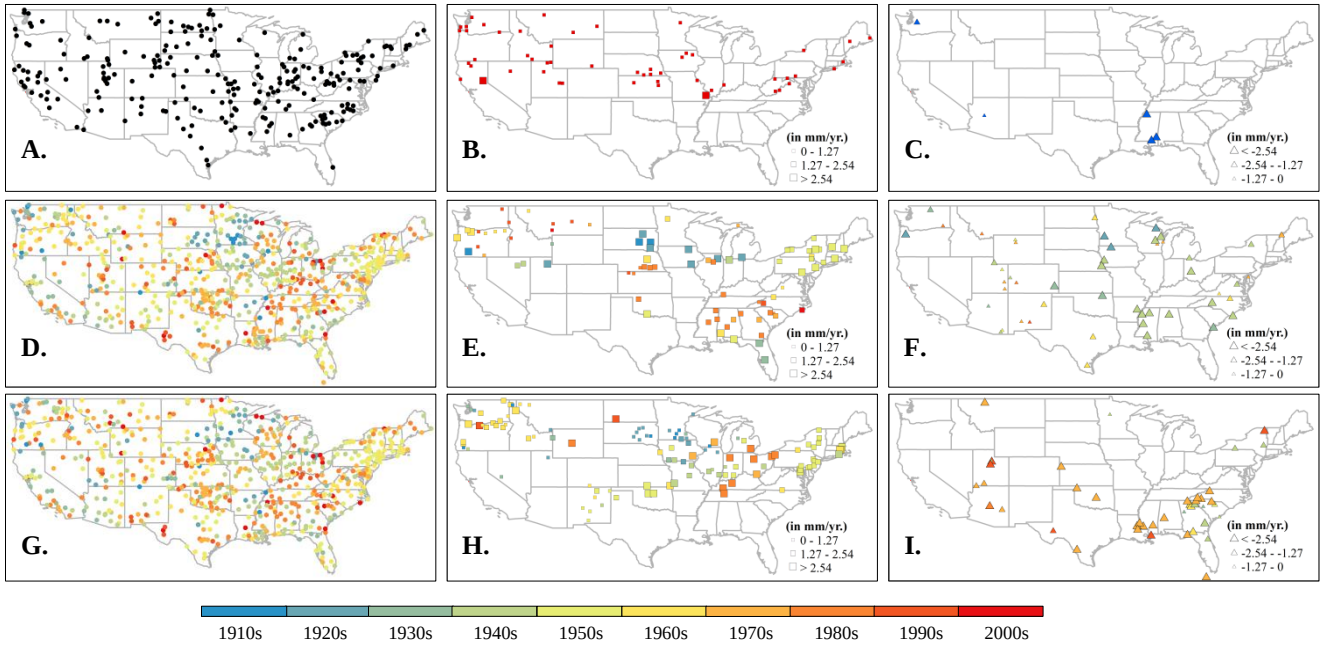


Figure A.6 (Cont.)

14) Spring total precipitation



15) Summer total precipitation

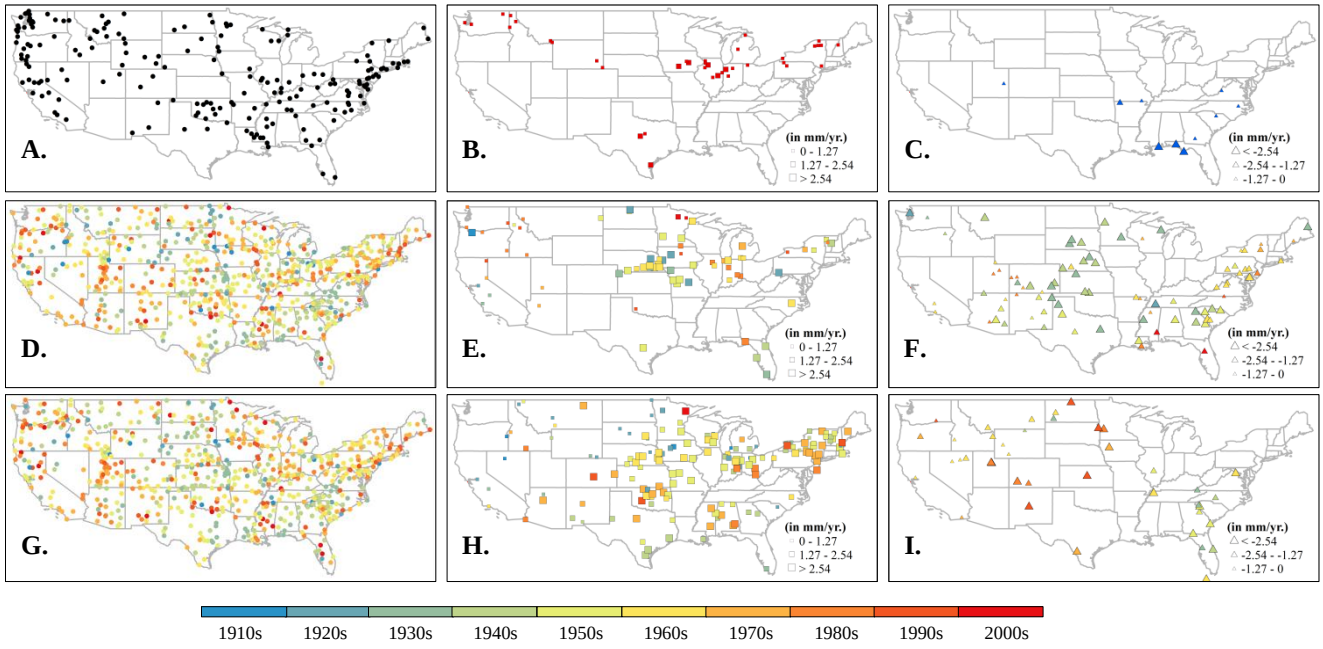


Figure A.6 (Cont.)

16) Fall total precipitation

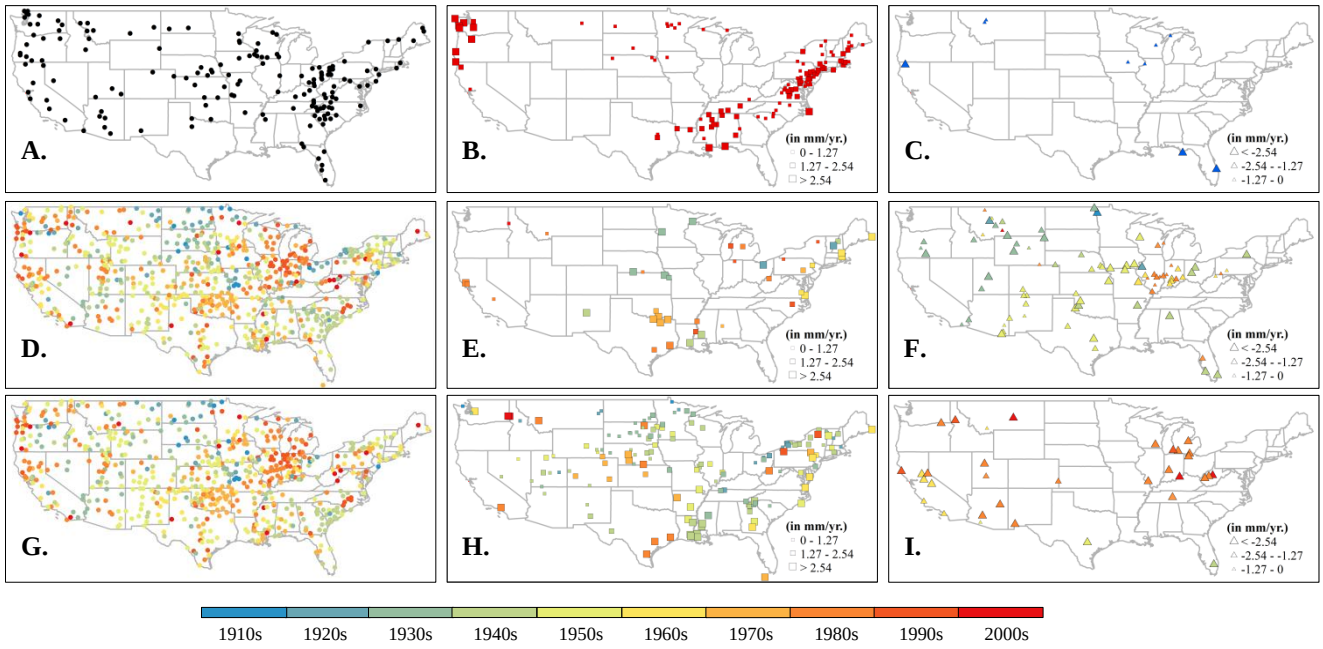
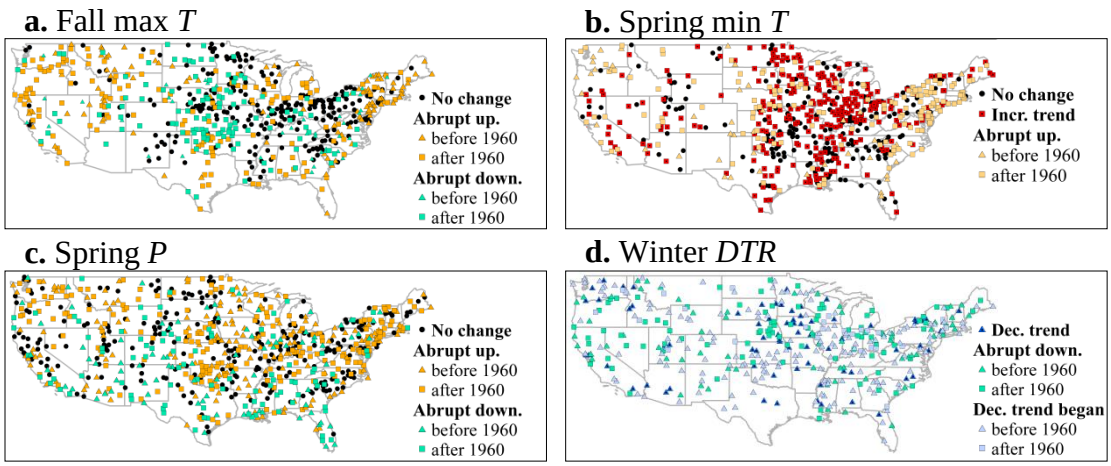


Figure A.6 (Cont.)



*In the above figure, abrupt change pattern means the time series only contain an abrupt change, which is between two states with constant means.

No change		Abrupt down.		Dec. trend	
Abrupt up.		Incr. trend		Dec. trend began	

Figure A.7 (**a**) Spatial distribution of "No change," "Abrupt upward change," and "Abrupt downward change" of fall mean maximum temperature; (**b**) Spatial distribution of "No change," "100-year increasing trend," and "Abrupt upward change" for spring mean minimum temperature; (**c**) Spatial distribution of "No change," "Abrupt upward change," and "Abrupt downward change" for spring total precipitation; (**d**) Spatial distribution of "100-year decreasing trend," "Abrupt downward change," and "Decreasing trend in later years" for winter mean DTR. (Schematic diagram for the mentioned change patterns are shown in the table.)

Appendix B: Supporting Information for Chapter 3

Supporting Information includes:

B.1: The process of generating de-changed climate data

B.2: Additional results

B.3: References

B.1 The process of generating de-changed climate data

The process starts with the introduction of the piecewise fit of a climate variable based on the original time series of the climate variable. The piecewise fit is conducted through the detection of the change pattern (Ge and Cai 2018) (see step 1 of figure S1) from the long-term mean of a climate variable in seasons during 1910-2010. The detection result shows whether there are abrupt change points; if there are, it also shows whether the trends in mean and the standard deviations of the linear parts segmented by the abrupt change points are positive, negative or zero-slope (Ge and Cai 2018). For example, figure S1 shows a time series with one abrupt downward shift preceded by a linear increasing trend before the change point and followed by a time period with a constant mean and linearly decreasing variance. More details of detecting abrupt shifts and trends in climate variables are shown in Chapter 2.

Following the detection of the abrupt shifts and the linear trends, these gradual and abrupt changes in mean and standard deviation are removed (de-changed) using the 1947 climate as a baseline (since 1947 is the beginning year with the crop yield data), as shown in step 2 of figure B.1. Equations (18) and (19) are used to generate the de-changed data (refer to step 3, de-changed data, of figure B.1).

The de-changed temperature is calculated by

$$\mathrm{Td}_{i,t} = \left(\mathrm{T}_{i,t} - \mathrm{T}^{*}_{i,t}\right) * \frac{\sigma^{*}_{i,t}}{\sigma_{i,1947}} + \mathrm{T}^{*}_{i,1947} \quad (18)$$

and the de-changed precipitation is calculated by

$$\mathrm{Pd}_{i,t} = \left(\mathrm{P}_{i,t} - \mathrm{P}^{*}_{i,t}\right) * \frac{\sigma^{*}_{i,t}}{\sigma_{i,1947}} + \mathrm{P}^{*}_{i,1947} \quad (19)$$

Where, $T^*_{i,t}(P^*_{i,t})$ is estimated temperature (precipitation) at county *i* at year *t*, from the piecewise fit during 1947-2010; $\sigma^*_{i,t}$ is estimated standard deviation for county *i* at year *t*, which is also based on the piecewise fit.

Note that the de-changing process is applied to maximum daily and minimum daily temperature, based on which, the average daily temperature and DTR are calculated.

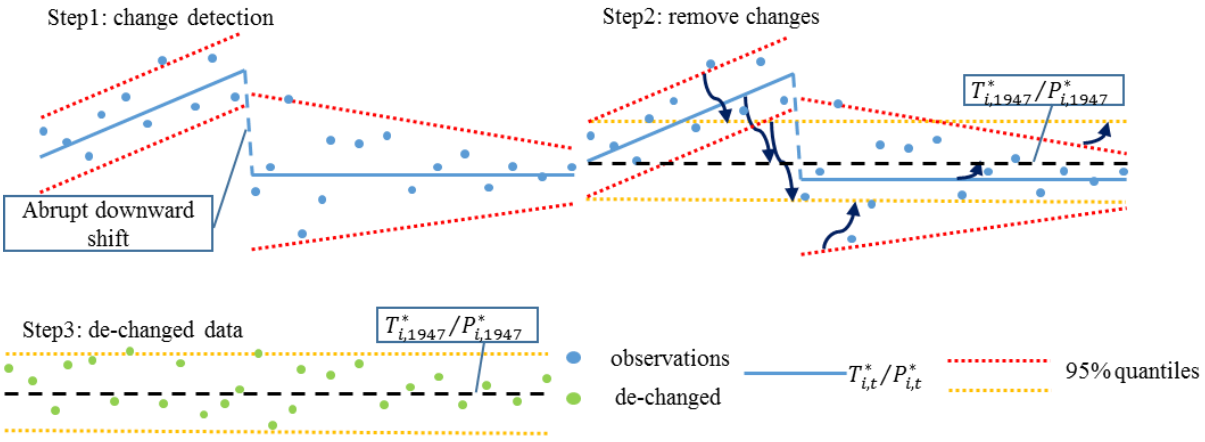


Figure B**.1.** A diagram showing the "de-changing" process. That is generating de-changed data by removing the changes identified from a climatic time series. The first step is to detect the gradual trends and abrupt shifts in mean and deviation; the second step is to remove the changes; in the third step, the de-changed data is generated.

B.2 Additional results

Table B.1 Percentages of counties in Nebraska exhibiting a upward parabolic or downward parabolic relationships between crop yield and precipitation or the various temperature variables in three regresson models (M2: P+Tmin, M3: P+Tavg, M4: P+DTR)

Irrigated	**Relationship**	**M2: P+Tmin**					**M3: P+Tavg**					**M4: P+DTR**				
Months (k=5, 6, 7, 8, 9)		**5**	**6**	**7**	**8**	**9**	**5**	**6**	**7**	**8**	**9**	**5**	**6**	**7**	**8**	**9**
Precipitation	Up-parabolic, $\alpha_{2ik} > 0$	55.2	50.6	17.2	43.7	62.1	48.3	51.7	24.1	51.7	56.3	55.2	37.9	23.0	52.9	51.7
	Down-parabolic, $\alpha_{2ik} < 0$	44.8	49.4	82.8	56.3	37.9	51.7	48.3	75.9	48.3	43.7	44.8	62.1	77.0	47.1	48.3
Temperature	Up-parabolic, $\beta_{2ik} > 0$	58.6	27.6	36.8	20.7	41.4	14.9	6.9	10.3	32.2	23.0	35.6	40.2	16.1	23.0	51.7
	Down-parabolic, $\beta_{2ik} < 0$	41.4	72.4	63.2	79.3	58.6	85.1	93.1	89.7	67.8	77.0	64.4	59.8	83.9	77.0	48.3
Rainfed	**Relationship**	**M2: P+Tmin**					**M3: P+Tavg**					**M4: P+DTR**				
Months (k=5, 6, 7, 8, 9)		**5**	**6**	**7**	**8**	**9**	**5**	**6**	**7**	**8**	**9**	**5**	**6**	**7**	**8**	**9**
Precipitation	Up-parabolic, $\alpha_{2ik} > 0$	34.5	34.5	5.7	19.5	71.3	46.0	42.5	16.1	26.4	65.5	42.5	50.6	8.0	43.7	64.4
	Down-parabolic, $\alpha_{2ik} < 0$	65.5	65.5	94.3	80.5	28.7	54.0	57.5	83.9	73.6	34.5	57.5	49.4	92.0	56.3	35.6
Temperature	Up-parabolic, $\beta_{2ik} > 0$	56.3	18.4	46.0	42.5	48.3	31.0	4.6	10.3	56.3	40.2	42.5	37.9	42.5	26.4	44.8
	Down-parabolic, $\beta_{2ik} < 0$	43.7	81.6	54.0	57.5	51.7	69.0	95.4	89.7	43.7	59.8	57.5	62.1	57.5	73.6	55.2

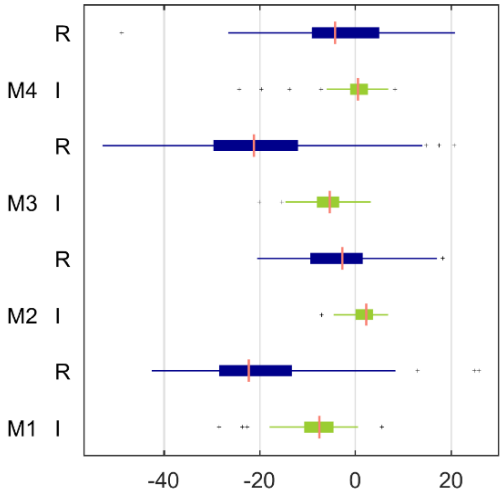


Figure B.2 Boxplots of parameter coefficients c_i in four regression models (M1: P+Tmax, M2: P+Tmin, M3: P+Tavg, M4: P+DTR) for irrigated and rainfed maize.

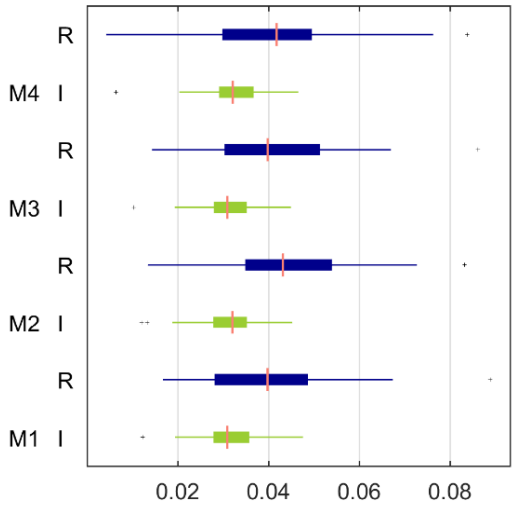


Figure B.3 Boxplots of parameter coefficients b_i in four regression models (M1: P+Tmax, M2: P+Tmin, M3: P+Tavg, M4: P+DTR) for irrigated and rainfed maize.

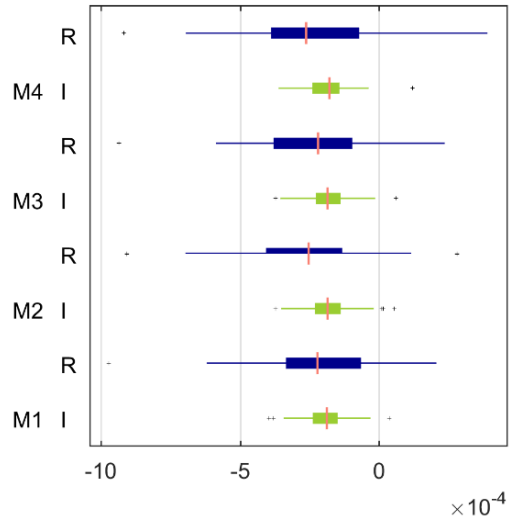


Figure B.4 Boxplots of parameter coefficients d_i in four regression models (M1: P+Tmax, M2: P+Tmin, M3: P+Tavg, M4: P+DTR) for irrigated and rainfed maize.

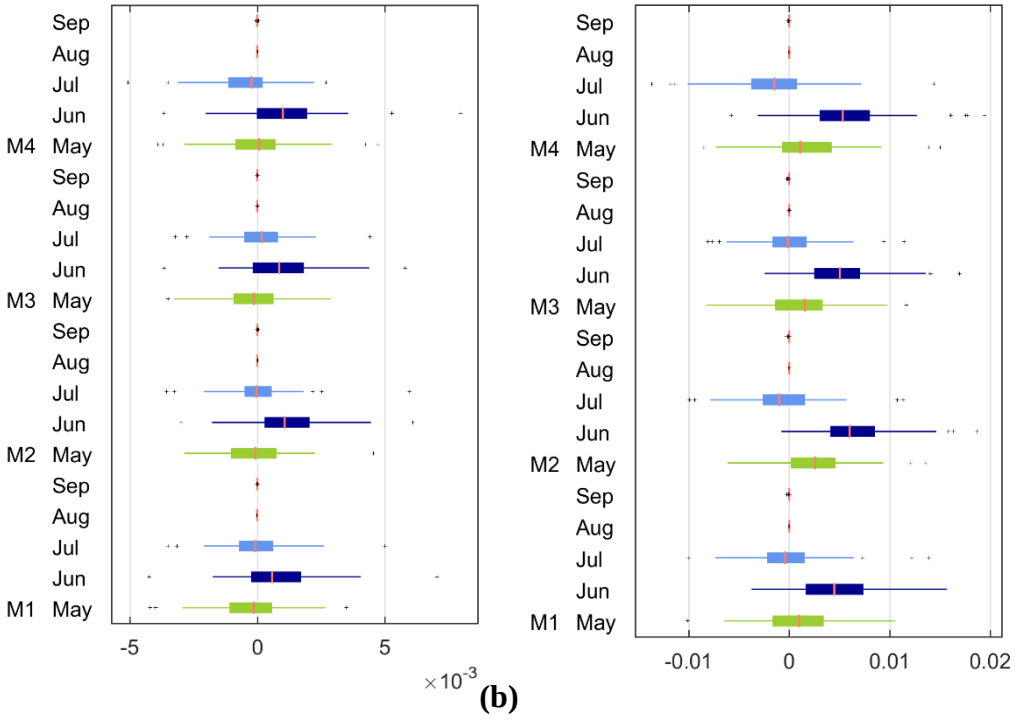
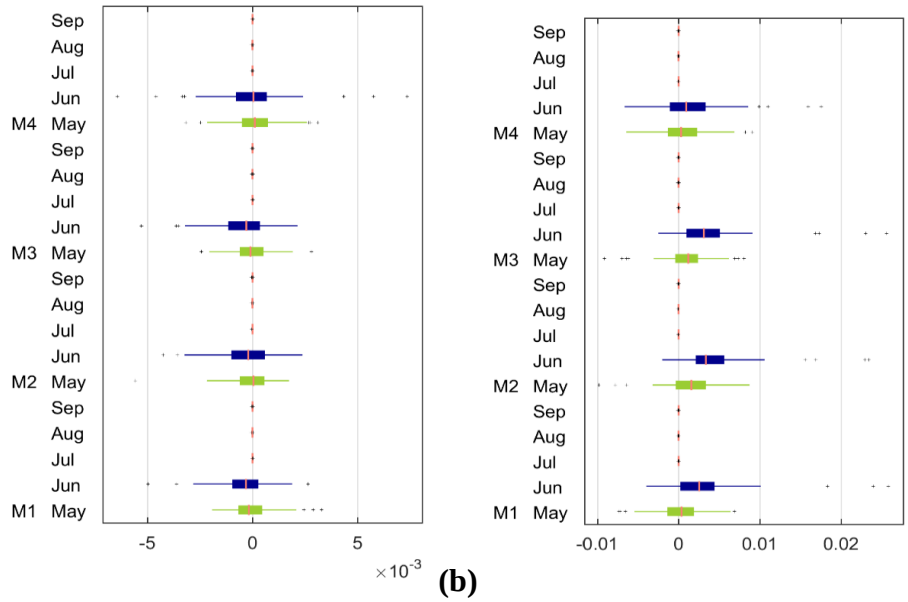
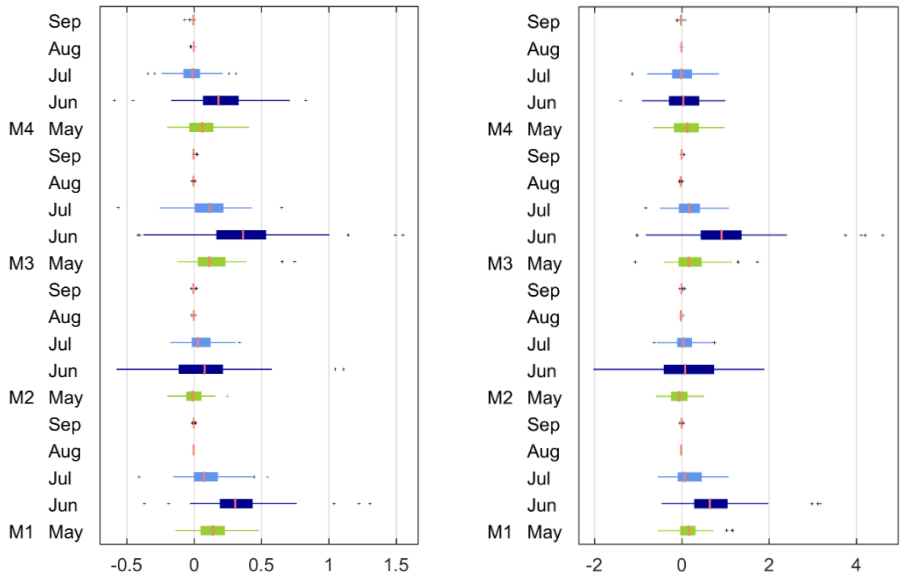


Figure B.5 Boxplots of parameter coefficients α_{1ik} in four regression models (M1: P+Tmax, M2: P+Tmin, M3: P+Tavg, M4: P+DTR) for (a) irrigated and (b) rainfed maize.



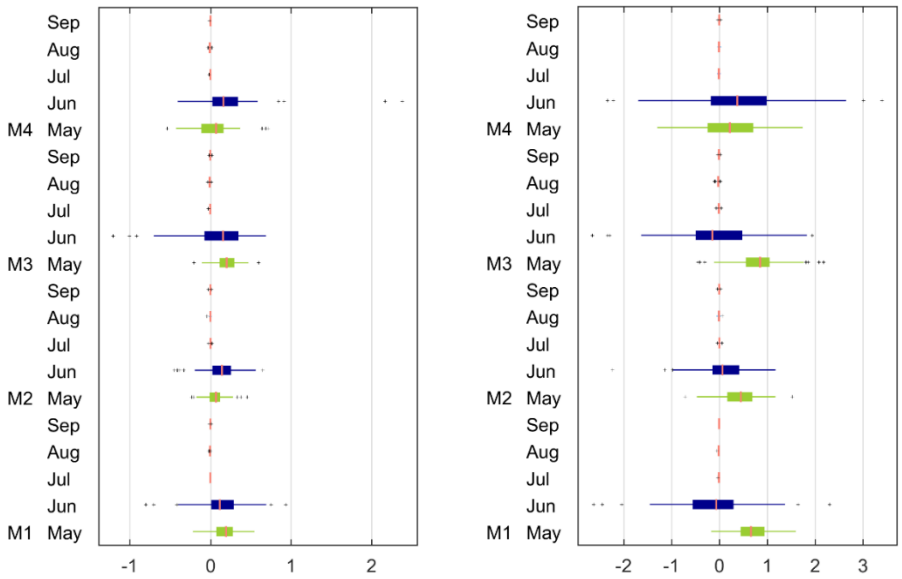
(a) **(b)**

Figure B.6 Boxplots of parameter coefficients α_{2ik} in four regression models (M1: P+Tmax, M2: P+Tmin, M3: P+Tavg, M4: P+DTR) for (a) irrigated and (b) rainfed maize.



(a) **(b)**

Figure B.7 Boxplots of parameter coefficients β_{1ik} in four regression models (M1: P+Tmax, M2: P+Tmin, M3: P+Tavg, M4: P+DTR) for (a) irrigated and (b) rainfed maize.



(a) **(b)**

Figure B.8 Boxplots of parameter coefficients β_{2ik} in four regression models (M1: P+Tmax, M2: P+Tmin, M3: P+Tavg, M4: P+DTR) for (a) irrigated and (b) rainfed maize.

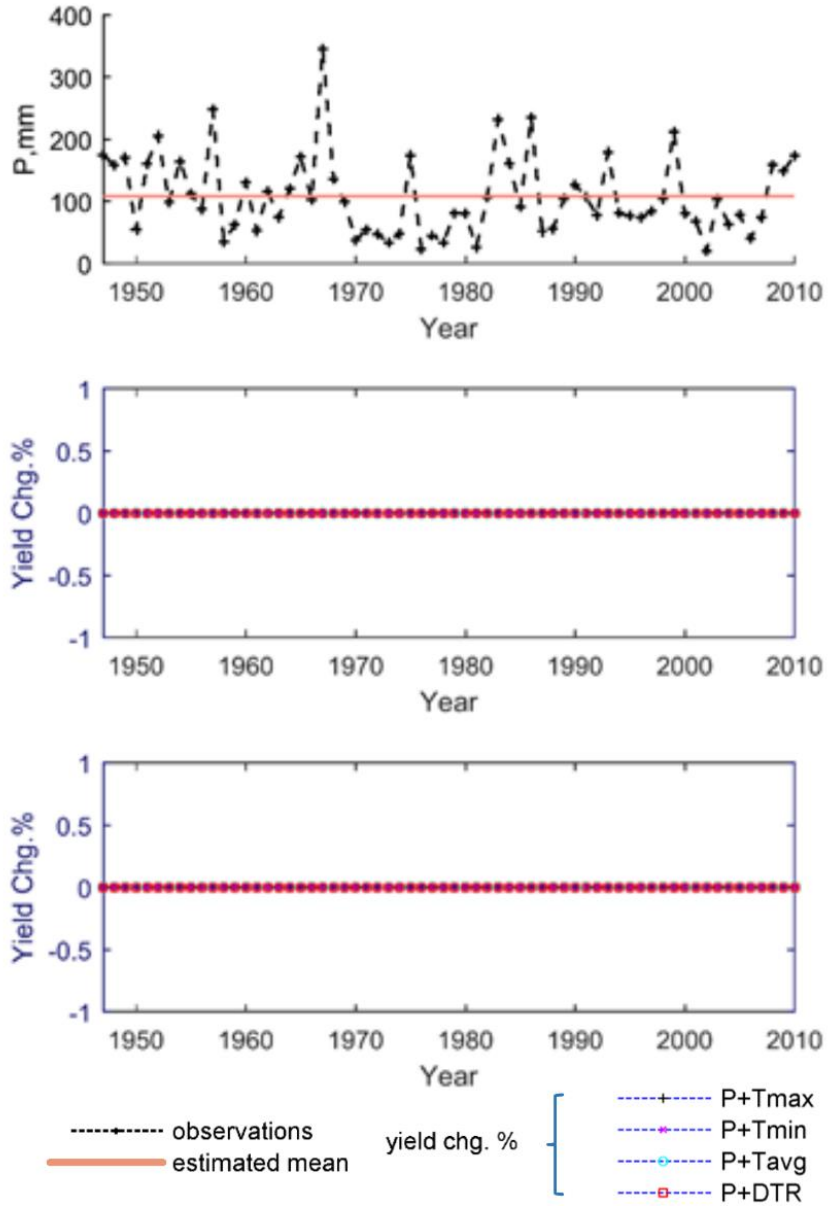
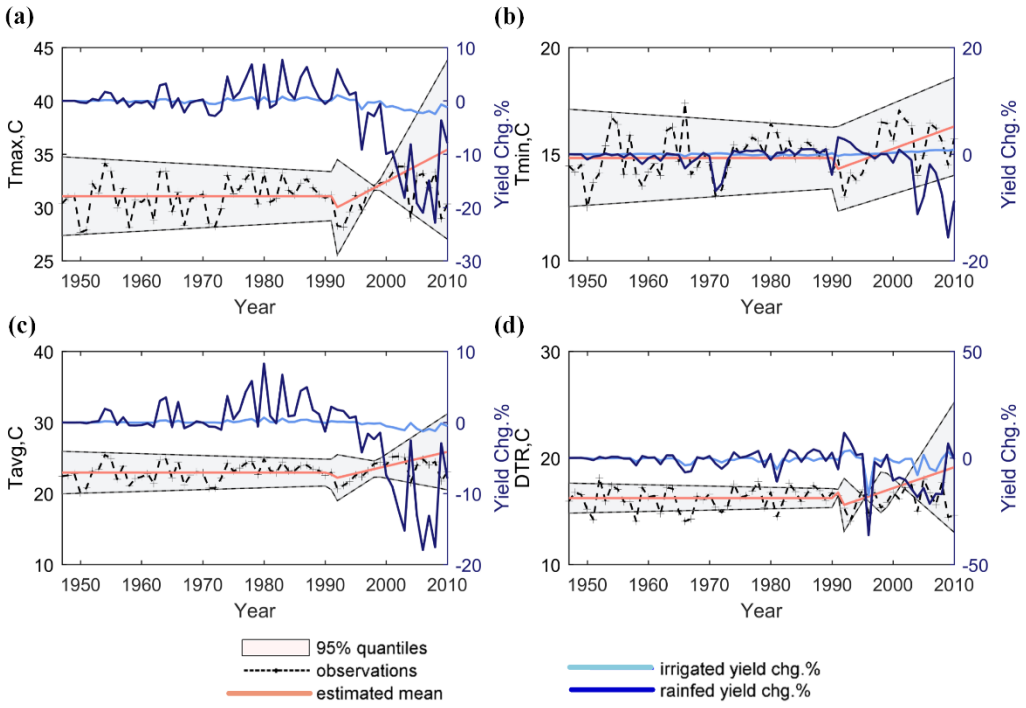


Figure B.9 Precipitation change over time and corresponding annual changes in irrigated and rainfed maize yield in Polk County when only linear trends in climate variables are assessed. (a) Changes in June precipitation; (b) changes in irrigated maize yield due to precipitation change, and (c) changes in rainfed maize yield due to precipitation change. Precipitation and one type of temperature are regressed with the logarithm value of crop yield in four models (i.e., P+Tmax, P+Tmin, P+Tavg and P+DTR)

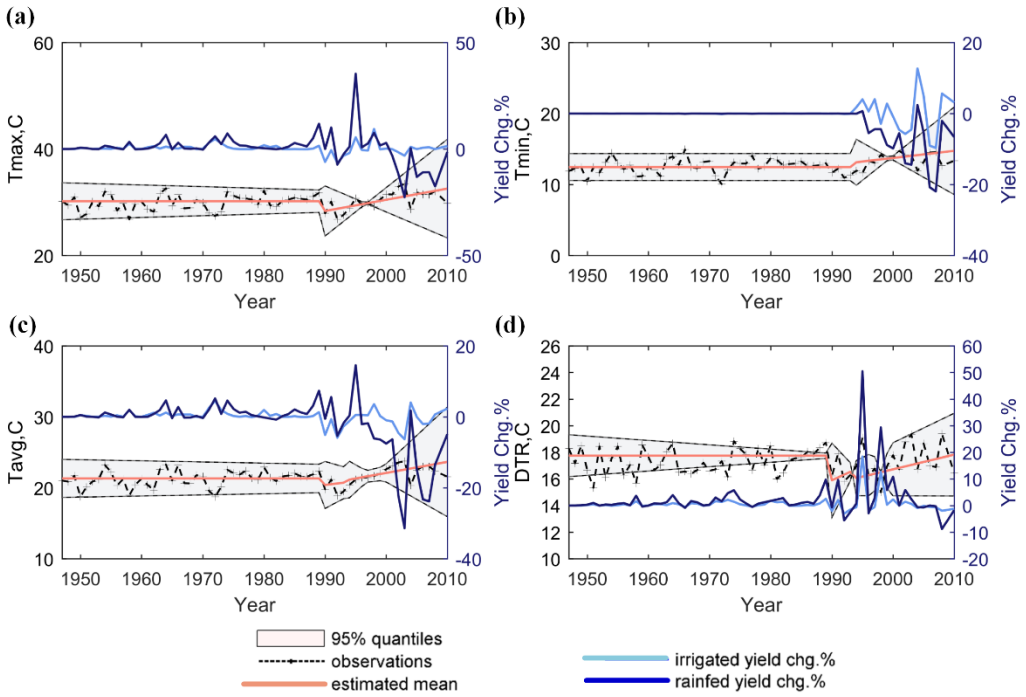


* 95% quantiles and estimated mean of DTR are those of daily maximum temperature minus the corresponding values of daily minimum temperature.

Contribution of temperature and precipitation change to rainfed maize yield change (%)

cumulative yield chg.% due to	P+Tmax	P+Tmin	P+Tavg	P+DTR
Temperature chg. (irrigated)	-0.55	0.16	-0.23	-1.5
Temperature chg. (rainfed)	-4.4	-2.7	-3.3	-4.4

Figure B.10 Temperature variable change over time and corresponding annual changes in irrigated and rainfed maize yield in Chase County. The changes in July mean temperature variables (a. *Tmax*, b. *Tmin*, c. *Tavg* and d**.** *DTR*) and how they impacted irrigated and rainfed maize yield. Precipitation and one type of temperature are regressed with log yield with each of the four models.



* 95% quantiles and estimated mean of DTR are those of daily maximum temperature minus the corresponding values of daily minimum temperature.

Contribution of temperature and precipitation change to rainfed maize yield change (%)

Cumulative yield chg.% due to	P+Tmax	P+Tmin	P+Tavg	P+DTR
Temperature chg. (irrigated)	0.17	-0.03	0.38	0.95
Temperature chg. (rainfed)	-2.8	-6.6	-7.7	2.5

Figure B.11 Temperature variable change over time and corresponding annual changes in irrigated and rainfed maize yield in Banner County. The changes in July mean temperature variables (a. *Tmax*, b. *Tmin*, c. *Tavg* and d**.** *DTR*) and how they impacted irrigated and rainfed maize yield. Precipitation and one type of temperature are regressed with log yield with each of the four models.

Finally, the normality assumption and the autocorrelation of regression model residuals can be concerns for the modeling exercise. We used the Kolmogorov-Smirnov method for a test. We found that the residuals of the models over 95% counties for irrigated maize and over 90% counties for rainfed maize were normally distributed. We also plotted the autocorrelation figure

(ACF) for each county. As an example, below in Figure B.12 is an ACF plot for rainfed maize in a representative county (ACF plots in most counties are similar to this one). No significant autocorrelation in any lag or trend in the regression residuals.

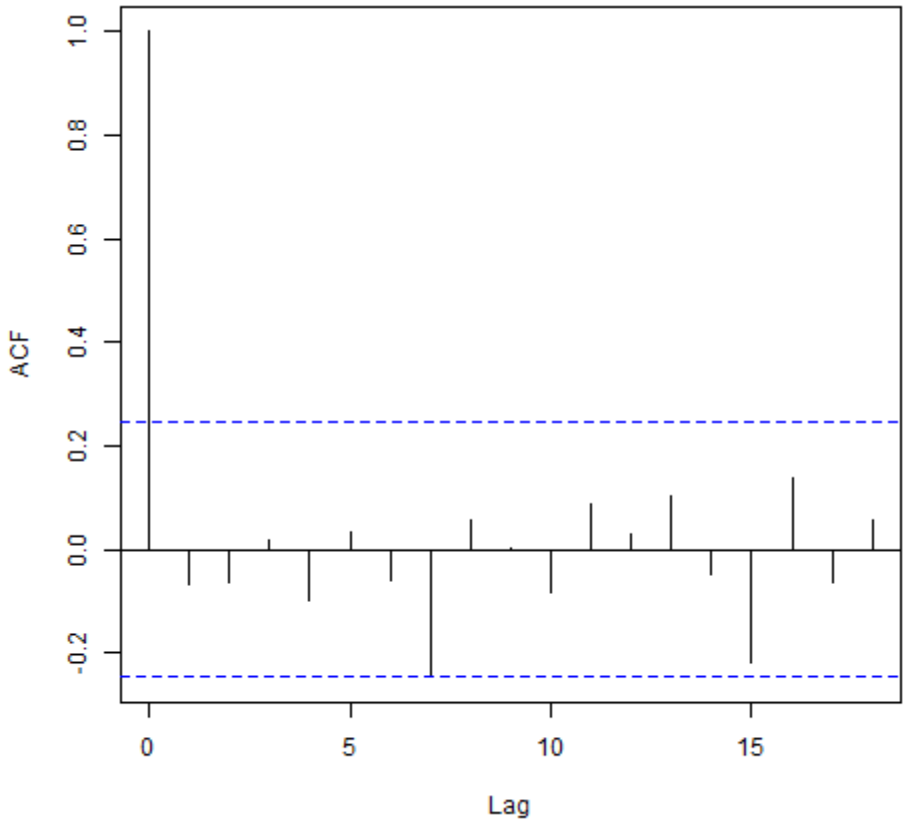


Figure B.12 Autocorrelation figure (ACF) for rainfed maize in a representative county

B.3 References

Chernoff, H., and Zacks, S. (1964), Estimating the Current Mean of a Normal Distribution which is Subjected to Changes in Time, *The Annals of Mathematical Statistics 35,* 999-1018.

Chu, P. S., and Zhao, X. (2011), Bayesian analysis for extreme climatic events: A review, *Atmospheric Research 102***,** 243-262.

Fong, D. K. H., and DeSarbo, W. S. (2007), A Bayesian methodology for simultaneously detecting and estimating regime change points and variable selection in multiple regression models for marketing research, *Qme-Quantitative Marketing and Economics 5*, 427-453.

Ge, Y. (2012), Detecting climate change and its impacts on crop yield in the continental United States *Master Thesis* Civil and Environmental Eng. University of Illinois at Urbana-Champaign

Ge, Y. and Cai, X. (2018), Abrupt shifts and gradual trends of temperature and precipitation of the past century in the continental United States, *under revision*

Hannart, A. and Naveau, P. (2009), Bayesian multiple change points and segmentation: Application to homogenization of climatic series, *Water Resources Research 45*, W10444.

Kim, J., and Cheon, S. (2010), Bayesian multiple change-point estimation with annealing stochastic approximation Monte Carlo, *Computational Statistics 25,* 215-239.

Lai, T. L. and Xing, H. P. (2011), A Simple Bayesian Approach to Multiple Change-Points, *Statistica Sinica 21,* 539-569

Lin, J. G., Chen, J., Li, Y. (2011), Bayesian Analysis of Student t Linear Regression with Unknown Change-Point and Application to Stock Data Analysis, *Computational Economics 40,* 203-217.

Pettitt, A. N. (1979) A non-parametric approach to the change point problem, *Appl. Statist. 28*, 126-135.

Ray, B. K. and Tsay, R. S. (2002), Bayesian methods for change-point detection in long-range dependent processes, *Journal of Time Series Analysis 23*, 687-705.

Rigaill, G., Lebarbier, E., Robin, S. (2012), Exact posterior distributions and model selection criteria for multiple change-point detection problems, *Statistics and Computing 22*, 917-929.

Rouge, C., Ge, Y., Cai, X. (2013), Detecting gradual and abrupt changes in hydrological records, *Advances in Water Resources 53*, 33-44.

Schutz, N., and Holschneider, M. (2011), Detection of trend changes in time series using Bayesian inference, *Physical Review E 84.*

Smith, A. F. M. (1975), A Bayesian Approach to Inference about a Change-Point in a Sequence of Random Variables, *Biometrika 62*, 407-416.

Tai, Y. C., Kvale, M. N., Witte, J. S. (2010), Segmentation and Estimation for SNP Microarrays: A Bayesian Multiple Change-Point Approach, *Biometrics 66*, 675-683.